
 RELEASE 3.1A JOHN F. COLLINS, BIOCOMPUTING RESEARCH UNIT.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSPRCH_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Sep 2 11:18:27 1999; MasPar time 8.07 Seconds
 60.852 Million cell updates/sec

 Output not generated.

Title: >US-08-599-226-3
 Description: (1-9) from US08599226.pep
 Perfect Score: 66
 Sequence: 1 QRYNRPYX 9

Scoring table: PAM 150
 Gap 15

Searched: 179066 segs, 54579741 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

sptrembl.9
 1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
 5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
 9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
 13:sp_vertebrate 14:sp_virus

Statistics: Mean 22.838; Variance 29.153; scale 0.783

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	57	86.4	560	5	044626	KIID12.1 PROTEIN.	3.24e+01
2	54	81.8	248	5	027562	PRT.	1.40e+00
3	54	81.8	260	13	042265	20S PROTEASOME SUBUNIT	1.40e+00
4	54	81.8	448	5	002441	GDP-DISSOCIATION INHIB	1.40e+00
5	54	81.8	1277	5	017517	ZC132.5 PROTEIN.	1.40e+00
6	53	80.3	790	2	055956	ABC TRANSPORTER.	2.27e+00
7	52	78.8	90	5	062346	RIL.2 PROTEIN.	3.64e+00
8	52	78.8	494	5	P92024	MEX-1.	3.64e+00
9	52	78.8	509	5	021017	COSMID F58H12.	3.64e+00
10	51	77.3	170	2	048279	SIMILIARTY WITH HEAD C	5.81e+00
11	50	75.8	178	4	092592	CD89_L10.	9.23e+00
12	50	75.8	201	4	099935	BASIC PROLINE-RICH PRO	9.23e+00
13	50	75.8	332	10	082438	ANTIFREEZE PROTEIN.	9.23e+00
14	50	75.8	450	2	050511	ZINC PROTEASE.	9.23e+00
15	50	75.8	521	2	044677	NEUTRAL PROTEASE.	9.23e+00
16	50	75.8	3082	14	036979	POLYPROTEIN.	9.23e+00
17	49	74.2	641	5	018586	COSMID C43H6	1.46e+01
18	48	72.7	43	10	049246	REVERSE TRANSCRIPTASE	2.28e+01
19	48	72.7	44	10	049247	REVERSE TRANSCRIPTASE	2.28e+01
20	48	72.7	192	5	019008	COSMID D2096.	2.28e+01

RESULT	1	PRELIMINARY;	PRT;	560 AA.	ALIGNMENTS
AC	044626;				
ID	044626;				
DT	01-JUN-1998 (TREMBLREL. 06, CREATED)				
DT	01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)				
DT	01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)				
DE	KIID12.1 PROTEIN.				
GN	KIID12.1				
OS	CAENORHABDITIS ELEGANS.				
OC	EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;				
OC	RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BRISTOL N2;				
RX	MEDLINE: 94150718.				
RA	WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,				
RA	BONEFIELD J., BIRTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,				
RA	CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,				
RA	GARDNER A., GREEN P., HAWKINS J., HILLIER L., JER M., JOHNSTON L.,				
RA	JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,				
RA	LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,				
RA	PARSONS J., PERCY C., RIKEN L., ROOPRA A., SAUNDERS D., SHANKEN R.,				
RA	SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SUJSTON J.,				
RA	THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,				
RA	WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,				
RA	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.				
RA	elegans.";				
RL	NATURE 368:32-38(1994).				
RP	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BRISTOL N2;				
RC	HENKHAUS J., WOHLDMANN P., GILLAM B.,				
RL	SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.				
RP	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BRISTOL N2;				
RC	WATERSTON R.,				
RL	SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.				
DR	EMBL: AF039047; G2736445; -				
DR	PROSITE: PS00028; ZINC_FINGER_C2H2.1				
KW	ZINC-FINGER; METAL-BINDING; DNA-BINDING.				
SO	SEQUENCE 560 AA; 61136 MW; 5906EBYD CRC32;				
Query Match	86.4%;	Score 57;	DB 5;	Length 560;	

Best Local Similarity 75.0%; Pred. No. 3,24e-01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 81 ORYKRAPY 88
|:|:|:|
QY 1 ORYKRAPY 8

RESULT 2
ID 027562 PRELIMINARY; PRT: 248 AA.
AC 027562;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRTC.
GN PRTC.
OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
OC EUKARYOTA; DICTYOSTELIIDA; DICTYOSTELIUM.
[1]
SEQUENCE FROM N.A.
SHAULSKY G., ESCALANTE R., LOOMIS W.F.;
SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S25A; ALSO KNOWN AS THE
CC PROTEASOME A-TYPE FAMILY.
CC -1- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
CC WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH
CC ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT
CC NEUTRAL OR SLIGHTLY BASIC PH.
CC THE PROTEASOME HAS AN ATP-DEPENDENT PROTEOLYTIC ACTIVITY.
CC THE COMPONENT C3 MAY HAVE A POTENTIAL REGULATORY EFFECT ON
CC ANOTHER COMPONENT(S) OF THE PROTEASOME COMPLEX THROUGH TYROSINE
CC PHOSPHORYLATION.
CC -1- PATHWAY: IS INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL
CC PROTEOLYTIC PATHWAY.
CC -1- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
CC SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE.
CC -1- SUBCELLULAR LOCATION: PROTEASOMES ARE FOUND IN THE CYTOPLASM AND
CC ALSO IN THE NUCLEUS.
DR EMBL; U60168; G1405364; -.
DR PROSITE; PS00388; PROTEASOME_A; 1.
DR PFAM; PF00227; proteasome; 1.
KW PROTEASOME; HYDROLASE; PROTEASE.
SQ SEQUENCE 248 AA; 2798 MW; BEE84ACC CRC32;

Query Match 81.8%; Score 54; DB 5; Length 248;
Best Local Similarity 75.0%; Pred. No. 1.40e+00;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

121 ORYKRAPY 128
|:|:|:|
QY 1 ORYKRAPY 8

RESULT 3
ID 042265 PRELIMINARY; PRT: 260 AA.
AC 042265;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE 205 PROTEASOME SUBUNIT C2.
GN CC2.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
CC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
[1]
SEQUENCE FROM N.A.
SINGH I., WAGNER B.J.;
SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF027978; G3136063; -.
DR PFAM; PF00227; proteasome; 1.
KW PROTEASOME.
SQ SEQUENCE 260 AA; 28925 MW; EF38499F CRC32;

Query Match 81.8%; Score 54; DB 13; Length 260;
Best Local Similarity 75.0%; Pred. No. 1.40e+00;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 120 ORYKRAPY 127
|:|:|:|
QY 1 ORYKRAPY 8

RESULT 4
ID 002441 PRELIMINARY; PRT: 448 AA.
AC 002441;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE GDI-DISSOCIATION INHIBITOR.
GN GDI.
OS GEODIA CIDONITUM (SPONGE).
OC EUKARYOTA; METAZOA; PORIFERA; DEMOSPONGIAE; TETRACITINOMORPHA;
CC ASTROPHORIDA; GEODIIDA; GEODIA.
[1]
SEQUENCE FROM N.A.
KRASKO A., SCHEFFER U., KOZIOL C., PANCER Z., BATEL R., BADRIA F.A.,
RA MUELLER W.E.G.;
RA AQUATIC TOXICOL. 37:157-168(1997).
DR EMBL; X94983; E218570; -.
DR PFAM; PF00996; GDI; 1.
SQ SEQUENCE 448 AA; 50215 MW; 848511EC CRC32;

Query Match 81.8%; Score 54; DB 5; Length 448;
Best Local Similarity 62.5%; Pred. No. 1.40e+00;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 217 ORYKRAPY 224
|:|:|:|
QY 1 ORYKRAPY 8

RESULT 5
ID 017517 PRELIMINARY; PRT: 1277 AA.
AC 017517;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ZC132.5 PROTEIN.
GN ZC132.5.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
CC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
[1]
SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAYELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLER L., JIR R., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LAURELLE P.,
RA LIGHTING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCT C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL NATURE 368:32-38(1994).
[2]
SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA BRADSHAW H., DEVLIN K.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
[3]
SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2:
RA WATERSTON R.:
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF014939; G2275629; -
SQ SEQUENCE 1277 AA; 146496 MW; 8A6D8899 CRC32;

Query Match 81.8%; Score 54; DB 5; Length 1277;
Best Local Similarity 100.0%; Pred. No. 1.40e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 877 QRYRRAP 883
QY 1 QRYRRAP 7

RESULT 6 PRELIMINARY; PRT; 790 AA.
AC 055956;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
ABC TRANSPORTER.
SYNECHOCYSTIS SP. (STRAIN PCC 6803).
OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PCC6803;
RX MEDLINE: 96127529.
RA KANEKO T., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,
RA SUGIURA M., TABATA S.,
RA "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. sequence features in the 1mb
RT region from map positions 648 to 928 of the genome.";
RL DNA RES. 2:153-166(1995).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-PCC6803;
RX MEDLINE: 97061201.
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSOUCHI T., MATSUO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
RA SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
RA TABATA S.,
RA "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. PCC6803. II. Sequence determination of the entire
RT genome and assignment of potential protein-coding regions.";
RL DNA RES. 3:109-136(1996).
DR EMBL: D64005; D1011375; -
PFAM: PF00005; ABC_tran; 1.
PFAM: PF00498; FHA; 2.
SQ SEQUENCE 790 AA; 87656 MW; 436D8E61 CRC32;

Query Match 80.3%; Score 53; DB 2; Length 790;
Best Local Similarity 62.5%; Pred. No. 2.27e+00;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 475 QRYRSRSY 482
QY 1 QRYRRAP 8

RESULT 7 PRELIMINARY; PRT; 90 AA.
AC 062346;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE R11.2 PROTEIN.
GN R11.2
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITIA; RHABDITIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]

RP SEQUENCE FROM N.A.
RA MCMURRAY A.:
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARONER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAM J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RT NATURE 368:32-38(1994).
DR EMBL: 281577; E1348694; -
SQ SEQUENCE 90 AA; 10937 MW; 80B83F8F CRC32;

Query Match 78.8%; Score 52; DB 5; Length 90;
Best Local Similarity 75.0%; Pred. No. 3.64e+00;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 21 QRYDRATY 28
QY 1 QRYRRAP 8

RESULT 8 PRELIMINARY; PRT; 494 AA.
AC P92024;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MEX-1.
GN MEX-1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITIA; RHABDITIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NZ;
RX MEDLINE: 97195715.
RA GUEDES S., PRIESS J.R.;
RT "The C. elegans MEX-1 protein is present in germline blastomeres and
RL development 124:731-739(1997).
DR EMBL: U81043; G1899062; -
PFAM: PF00642; Zf-CCCH; 2.
SQ SEQUENCE 494 AA; 56608 MW; 716506D2 CRC32;

Query Match 78.8%; Score 52; DB 5; Length 494;
Best Local Similarity 62.5%; Pred. No. 3.64e+00;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 263 QRYRRPPF 270
QY 1 QRYRRAP 8

RESULT 9 PRELIMINARY; PRT; 509 AA.
AC Q21017;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE COSMID F58H12.
GN F58H12.1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC

OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 [1]
 RN SEQUENCE FROM N.A.
 RA MEDLINE: 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BOFIELD J., BURTON J., CONNELL M., COPESEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN T., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHONKKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULLSTON J.,
 RA THERYER-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHIDMAN P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL NATURE 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RN MILLER N.;
 RN SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA WATERSTON R.;
 RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U0416; G1063490; -.
 SQ SEQUENCE 509 AA; 55953 MW; D0EC1387 CRC32;
 Query Match 78.8%; Score 52; DB 5; Length 509;
 Best Local Similarity 71.4%; Pred. No. 3.64e+00;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Db 268 RYRNPY 274
 QY 2 RYRNPY 8
 ID 048279 PRELIMINARY; PRT; 170 AA.
 AC 048279;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
 DE SIMILARITY WITH HEAD COMPLETION/STABILIZATION PROTEIN.
 OS HAEMOPHILUS SOMNUS.
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;
 OC HAEMOPHILUS.
 OC [1]
 RN SEQUENCE FROM N.A.
 RN STRAIN-HS25;
 RA POSTAROLLO R.A.;
 RL THESIS (1995), V. I. D. O., UNIVERSITY OF SASKATCHEWAN.
 DR EMBL: U28154; G915368; -.
 SQ SEQUENCE 170 AA; 19923 MW; 91428D5E CRC32;
 Query Match 77.3%; Score 51; DB 2; Length 170;
 Best Local Similarity 75.0%; Pred. No. 5.81e+00;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Db 107 QRYKRVY 114
 QY 1 QRYKRVY 8
 ID 092592 PRELIMINARY; PRT; 178 AA.
 AC 092592;
 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
 DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
 DE C889.L10.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;

OC CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RN SEQUENCE FROM N.A.
 RA TOYABE S., KUMANO Y., TAKEDA K., UCHIYAMA M., ABO T.;
 RT "Alternatively spliced forms of monocytic Iga Fc receptors in patients
 RT with Iga nephropathy.";
 RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: D87861; D101171; -.
 SQ SEQUENCE 178 AA; 19593 MW; EC5B1F1C CRC32;
 Query Match 75.8%; Score 50; DB 4; Length 178;
 Best Local Similarity 71.4%; Pred. No. 9.23e+00;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Db 87 WYNSRPY 93
 QY 2 WYNSRPY 8
 ID 099935 PRELIMINARY; PRT; 201 AA.
 AC 099935;
 DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
 DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE BASIC PROLINE-RICH PROTEIN.
 GN BLP.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
 OC CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RN SEQUENCE FROM N.A.
 RA MEDLINE: 96309097.
 RA DICKINSON D.P., THIESSE M.;
 RT "CDNA cloning of an abundant human lacrimal gland mRNA encoding a
 RT novel tear protein.";
 RL CURR. EYE RES. 15:377-386(1996).
 DR EMBL: S83198; G1836022; -.
 SQ SEQUENCE 201 AA; 22870 MW; 650A8F70 CRC32;
 Query Match 75.8%; Score 50; DB 4; Length 201;
 Best Local Similarity 62.5%; Pred. No. 9.23e+00;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Db 22 QRFSSRPY 29
 QY 1 QRYNRPY 8
 ID 082438 PRELIMINARY; PRT; 332 AA.
 AC 082438;
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE ANTIFREEZE PROTEIN.
 GN AFP.
 OS DAUCUS CAROTA (CARROT).
 OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
 OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;
 OC ASTERIDAE; ANALES; APIACEAE; DAUCUS.
 RN [1]
 RN SEQUENCE FROM N.A.
 RN STRAIN-CV, AUTUMN KING, TISSUE-TAP ROOT;
 RX MEDLINE: 98429644.
 RA MORRILL D., ELIAS L., ASHFORD D., SMALLWOOD M., SIDEBOTTOM C.,
 RA LITTLE P., TELFORD J., HOLT C., BOWLES D.;
 RT "A carrot leucine-rich-repeat protein that inhibits ice
 RT recrystallization.";
 RL SCIENCE 282:115-117(1998).
 DR EMBL: AF054480; G3702803; -.
 SQ SEQUENCE 332 AA; 36845 MW; 67194791 CRC32;

Query Match 75.8%; Score 50; DB 10; Length 332;
 Best Local Similarity 50.0%; Pred. No. 9.23e+00;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 DB 311 QRYDRTAY 318
 ||:|:|:|
 QY 1 QRYNRAPY 8

RESULT 14
 ID 050511; PRELIMINARY; PRT; 450 AA.

AC 050511;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE ZINC PROTEASE.
 GN SC9810.04.

OS STREPTOMYCES COELICOLOR.

OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
 OC ACTINOMYCETALES; STREPTOMYCINAE; STREPTOMYCETACEAE; STREPTOMYCES.

[1]
 SEQUENCE FROM N.A.

STRAIN-A3(2);
 OLIVER K., HARRIS D.;

RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

[2]
 SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;

RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

[3]
 SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RX MEDLINE: 97000351.

RA REDENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,

RA KINASHI H., HOPWOOD D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL MOL. MICROBIOL. 21:77-96(1996).

DR EMBL: AL009204; E1202335; -.

KW PROTEASE.

SQ SEQUENCE 450 AA; 49045 MW; 3E30B71A CRC32;

Query Match 75.8%; Score 50; DB 2; Length 450;

Best Local Similarity 62.5%; Pred. No. 9.23e+00;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

DB 147 QRYDNVPY 154
 |||:|:|
 1 QRYNRAPY 8

RESULT 15
 ID Q44677; PRELIMINARY; PRT; 521 AA.

AC Q44677;

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)

DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE NEUTRAL PROTEASE.

OS BACILLUS AMYLOLIQUEFACIENS.

OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;

OC BACILLUS.

RN [1]

RP SEQUENCE FROM N.A.

RA SHIMADA H., HONJO M., MITA I., NAKAYAMA A., AKAOKA A., MANABE K.,

RA FURUTANI Y.;

RL J. BIOTECHNOL. 2:75-85(1985).

DR EMBL: M36723; G143353; -.

DR PRAM: P000099; zn-protease; 1.

KW PROTEASE.

SQ SEQUENCE 521 AA; 56725 MW; 64AFEE5F CRC32;

Query Match 75.8%; Score 50; DB 2; Length 521;

Best Local Similarity 62.5%; Pred. No. 9.23e+00;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 DB 308 QRYNRSY 315
 |:|:|:|:|
 QY 1 QRYNRAPY 8

Search completed: Thu Sep 2 11:18:44 1999
 Job time : 17 secs.

THIS PAGE BLANK (USPTO)

WISER (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSrch.PP protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Sep 2 11:21:51 1999; Maspar time 4.46 Seconds
146.856 Million cell updates/sec

Molecular output not generated.

Title: >US-08-599-226-4
Description: (1-12) from US08599226.Pep
Sequence: 66
1 VSYLSTASSLDX 12

Scoring table: PAM 150
Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

5:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human
5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.organelle
9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified
13:sp.vertebrate 14:sp.virus

Statistics: Mean 23.196; Variance 28.631; scale 0.810

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	55	83.3	649	3	PROTEIN COMPLEX ASSEMB	8.09e-01
2	51	77.3	236	3	CHROMOSOME XV READING	5.65e+00
3	51	77.3	3623	11	INTRINSIC FACTOR-B12 R	5.65e+00
4	50	75.8	293	2	CONSERVED HYPOHETICAL	9.06e+00
5	50	75.8	705	2	CADMIUM RESISTANCE PRO	9.06e+00
6	49	74.2	599	2	PHOSPHONOLPYRUVATE CA	1.44e+01
7	48	72.7	292	14	PUTATIVE RIBONUCLEASE	2.28e+01
8	48	72.7	467	4	KIAA0414.	2.28e+01
9	48	72.7	832	4	P/CAF.	2.28e+01
10	48	72.7	1245	5	F40E12.2 PROTEIN (FRAG	2.28e+01
11	47	71.2	127	14	NUCLEOPROTEIN (FRAGME	3.58e+01
12	47	71.2	149	11	MYELIN BASIC PROTEIN (3.58e+01
13	47	71.2	195	11	MYELIN BASIC PROTEIN (3.58e+01
14	47	71.2	197	4	GOLLI-MBP.	3.58e+01
15	47	71.2	239	2	SRMx PROTEIN.	3.58e+01
16	47	71.2	250	11	MYELIN BASIC PROTEIN (3.58e+01
17	47	71.2	255	10	AP2 DOMAIN CONTAINING	3.58e+01
18	47	71.2	303	14	SIMILAR TO PBCV-1 SERI	3.58e+01
19	47	71.2	304	4	GOLLI-MBP.	3.58e+01
20	47	71.2	458	4	KIAA0582 PROTEIN (FRAG	3.58e+01

21	47	71.2	959	5	018359	PUTATIVE KRUEPPEL TARG	3.58e+01
22	47	71.2 <th>1128</th> <th>1</th> <th>051009</th> <th>REP1 PROTEIN.</th> <th>3.58e+01</th>	1128	1	051009	REP1 PROTEIN.	3.58e+01
23	47	71.2 <th>1128</th> <th>1</th> <th>051999</th> <th>HYPOHETICAL. 14.7 KD P</th> <th>5.58e+01</th>	1128	1	051999	HYPOHETICAL. 14.7 KD P	5.58e+01
24	46	69.7 <th>133</th> <th>2</th> <th>084670</th> <th>COSMID F28B12.</th> <th>5.58e+01</th>	133	2	084670	COSMID F28B12.	5.58e+01
25	46	69.7 <th>258</th> <th>5</th> <th>019850</th> <th>HYPOHETICAL PROTEIN (</th> <th>5.58e+01</th>	258	5	019850	HYPOHETICAL PROTEIN (5.58e+01
26	46	69.7 <th>296</th> <th>14</th> <th>069118</th> <th>ANTHRANILATE SYNTHASE</th> <th>5.58e+01</th>	296	14	069118	ANTHRANILATE SYNTHASE	5.58e+01
27	46	69.7 <th>500</th> <th>2</th> <th>025869</th> <th>AMINO ACID PERMEASE</th> <th>5.58e+01</th>	500	2	025869	AMINO ACID PERMEASE	5.58e+01
28	46	69.7 <th>503</th> <th>5</th> <th>026594</th> <th>A-KINASE ANCHORING PRO</th> <th>5.58e+01</th>	503	5	026594	A-KINASE ANCHORING PRO	5.58e+01
29	46	69.7 <th>1022</th> <th>6</th> <th>028628</th> <th>RNA POLYMERASE I.</th> <th>5.58e+01</th>	1022	6	028628	RNA POLYMERASE I.	5.58e+01
30	46	69.7 <th>2910</th> <th>5</th> <th>026008</th> <th>GLYCOPROTEIN GJ.</th> <th>8.64e+01</th>	2910	5	026008	GLYCOPROTEIN GJ.	8.64e+01
31	45 <th>68.2</th> <th>121</th> <th>14</th> <th>090626</th> <th>GLYCOPROTEIN GJ.</th> <th>8.64e+01</th>	68.2	121	14	090626	GLYCOPROTEIN GJ.	8.64e+01
32	45 <th>68.2</th> <th>131</th> <th>14</th> <th>090638</th> <th>COSMID ZC64.</th> <th>8.64e+01</th>	68.2	131	14	090638	COSMID ZC64.	8.64e+01
33	45 <th>68.2</th> <th>164</th> <th>5</th> <th>023378</th> <th>261AA LONG HYPOHETICA</th> <th>8.64e+01</th>	68.2	164	5	023378	261AA LONG HYPOHETICA	8.64e+01
34	45 <th>68.2</th> <th>261</th> <th>1</th> <th>059473</th> <th>ATP SYNTHASE A CHAIN (</th> <th>8.64e+01</th>	68.2	261	1	059473	ATP SYNTHASE A CHAIN (8.64e+01
35	45 <th>68.2</th> <th>276</th> <th>8</th> <th>033658</th> <th>HYPOHETICAL 42.1 KD P</th> <th>8.64e+01</th>	68.2	276	8	033658	HYPOHETICAL 42.1 KD P	8.64e+01
36	45 <th>68.2</th> <th>370</th> <th>10</th> <th>065492</th> <th>HYPOHETICAL 41.9 KD P</th> <th>8.64e+01</th>	68.2	370	10	065492	HYPOHETICAL 41.9 KD P	8.64e+01
37	45 <th>68.2</th> <th>384</th> <th>2</th> <th>011197</th> <th>NWEG PRECURSOR (EC 3.2</th> <th>8.64e+01</th>	68.2	384	2	011197	NWEG PRECURSOR (EC 3.2	8.64e+01
38	45 <th>68.2</th> <th>448</th> <th>5</th> <th>077045</th> <th>HYPOHETICAL 52.8 KD P</th> <th>8.64e+01</th>	68.2	448	5	077045	HYPOHETICAL 52.8 KD P	8.64e+01
39	45 <th>68.2</th> <th>471</th> <th>10</th> <th>023552</th> <th>MERCURIC REDUCTASE.</th> <th>8.64e+01</th>	68.2	471	10	023552	MERCURIC REDUCTASE.	8.64e+01
40	45 <th>68.2</th> <th>546</th> <th>2</th> <th>069236</th> <th>CDCL6HS.</th> <th>8.64e+01</th>	68.2	546	2	069236	CDCL6HS.	8.64e+01
41	45 <th>68.2</th> <th>608</th> <th>10</th> <th>P92973</th> <th>MULTI-FUNCTIONAL ENZYM</th> <th>8.64e+01</th>	68.2	608	10	P92973	MULTI-FUNCTIONAL ENZYM	8.64e+01
42	45 <th>68.2</th> <th>619</th> <th>4</th> <th>013042</th> <th>TYROSINE KINASE RECEPT</th> <th>8.64e+01</th>	68.2	619	4	013042	TYROSINE KINASE RECEPT	8.64e+01
43	45 <th>68.2</th> <th>631</th> <th>2</th> <th>086216</th> <th></th> <th></th>	68.2	631	2	086216		
44	45 <th>68.2</th> <th>650</th> <th>2</th> <th>050417</th> <th></th> <th></th>	68.2	650	2	050417		
45	45 <th>68.2</th> <th>1369</th> <th>13</th> <th>P79950</th> <th></th> <th></th>	68.2	1369	13	P79950		

ALIGNMENTS

RESULT 1
ID 060167; PRELIMINARY; PRT; 649 AA.

AC 060167;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE PROTEIN COMPLEX ASSEMBLY PROTEIN.
GN SPIC19F8.03C.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHITASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA BECK A., REINHARDT R., WOOD V., RAJANDREAM M.A., BARRELL B.G.;
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AL023594; E1293401; -;
SQ SEQUENCE 649 AA; 72985 MW; 9C207DB2 CRC32;

Query Match 83.3%; Score 55; DB 3; Length 649;
Best Local Similarity 63.6%; Pred. No. 8.09e-01;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 233 INVSTARSLE 243
QY 1 VSYLSTASSLD 11
RESULT 2
ID 012282; PRELIMINARY; PRT; 236 AA.
AC 012282;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE CHROMOSOME XV READING FRAME ORF YOR214C.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RA BOYER J., FAIRHEAD C., GAILLON L., GALLISSON F., MICHAUX G.,
RA THIERRY A., DUJON B.;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

```

RN [2]
RP SEQUENCE FROM N.A.
RA MIPG;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-FY1679;
RA GALLISON F., DUJON B.;
RL SUBMITTED (OCT-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; 275122; E252081;
DR EMBL; X92441; G1050766;
SQ SEQUENCE 236 AA; 26156 MW; 9759C243 CRC32;

Query Match
Best Local Similarity 60.0%; Score 51; DB 3; Length 236;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 175 IYLSSTSL 184
:||||:|
1 VSYLSTASSL 10

RESULT 3
ID 070244 PRELIMINARY; PRT: 3623 AA.
AC 070244;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-JAN-1999 (TREMBLREL. 09, LAST SEQUENCE UPDATE)
DE 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DT INTRINSIC FACTOR-B12 RECEPTOR PRECURSOR.
GN CUBILIN.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIOGNATHI; MURIDAE; MORINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98148073.
RA MOESTRUP S.K., KOZYRAKI R., KRISTIANSEN M., KAYSEN J.H.,
RA RASMUSSEN H.H., BRAUT D., PONTILION F., GODA F.O., CHRISTENSEN E.I.,
RA HAMMOND T.G., VERROUST P.J.;
RT "The intrinsic factor-vitamin B12 receptor and target of teratogenic
RT antibodies is a megalin-binding peripheral membrane protein with
RT homology to developmental proteins."
RT J. Biol. Chem. 273:5235-5242(1998).
DR EMBL; AF022247; G3834380;
DR PROSITE; PS001010; ASX-HYDROXYL; 3.
DR SIGNAL; GLYCOPROTEIN; EGF-LIKE DOMAIN.
KW SIGNAL.
FT CHAIN 1 20 POTENTIAL.
FT SIGNAL 21 3623 INTRINSIC FACTOR-B12 RECEPTOR.
SQ SEQUENCE 3623 AA; 398981 MW; ADF804DC CRC32;

Query Match
Best Local Similarity 77.3%; Score 51; DB 11; Length 3623;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 697 IYLTQSDLD 707
:||||:|
1 VSYLSTASSL 11

RESULT 4
ID 051095 PRELIMINARY; PRT: 293 AA.
AC 051095;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CONSERVED HYPOTHETICAL PROTEIN.
GN BBO068.
OS BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).
OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / B31;

```

```

RX MEDLINE; 98065943.
RA FRASER C.M., CASSENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA LATHIGRA R., WHITE O., KETCHUM K.A., DOBSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERAVAGE A.R., OUACKENBUSH J., SALZBERG S., HANSON M.,
RA VAN VIGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTTERBACK T., WATTHEY L., McDONALD L., ARTTACH P., BOWMAN C.,
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi."
RT NATURE 390:580-586(1997).
DR EMBL; AE001120; G2687956;
DR TIGR; BB0068;
SQ SEQUENCE 293 AA; 33278 MW; 3FB9B9E2 CRC32;

Query Match
Best Local Similarity 75.8%; Score 50; DB 2; Length 293;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 198 AYLSSTNSLE 207
:||||:|
2 SYLSTASSLD 11

RESULT 5
ID P94888 PRELIMINARY; PRT: 705 AA.
AC P94888;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CADMIUM RESISTANCE PROTEIN.
GN CADA.
OS LACTOCOCCUS LACTIS (STREPTOCOCCUS LACTIS).
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE.
OC LACTOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-M71;
RA LIU C.O., CHIA G.-L., DUNN N.W.;
RT SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U78967; G1699049;
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
DR PFAM; PF00122; E1-E2_ATPase; 1.
DR PFAM; PF00403; HMA; 1.
KW PLASMIN; HYDROLASE; TRANSMEMBRANE; PHOSPHORYLATION; ATP-BINDING.
FT MOD_RES 398
FT SEQUENCE 705 AA; 76466 MW; 6648C08D CRC32;

Query Match
Best Local Similarity 75.8%; Score 50; DB 2; Length 705;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 422 IYLSITSSLE 432
:||||:|
1 VSYLSTASSLD 11

RESULT 6
ID 084716 PRELIMINARY; PRT: 599 AA.
AC 084716;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PHOSPHONOLPYRUVATE CARBOXYKINASE.
GN PCXA.
OS CHLAMYDIA TRACHOMATIS.
OC BACTERIA; CHLAMYDIALES; CHLAMYDIACEAE; CHLAMYDIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D/TW-3/CX.
RA STEPHENS R.S., KALMAN S., LAMMEL C.J., PAN J., MARATHE R., ARAVIND L.,

```

RA MITCHELL W.P., OLINGER L., TATISOV R.L., ZHAO Q., KOONIN E.V.,
RA DAVIS R.W.: "Genome Sequence of an Obligate Intracellular Pathogen of Humans:
RT Chlamydia trachomatis.";
RL SCIENCE 0:0-0(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STEPHAN-D'UM-3/CX;
RA STEPHANS R.S., KALMAN S., LAMMEL C.J., FAN J., MARACHE R., ARAVIND L.,
RA MITCHELL W.P., OLINGER L., TATISOV R.L., ZHAO Q., KOONIN E.V.,
RA DAVIS R.W.;
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DDBJ DATA BANKS.
RL EMBL; AE001341; G3329165; -
DR PROSITE; PS00505; BEPCCK_GRP; 1.
KW PRIVATE.
SQ SEQUENCE 599 AA; 66244 MW; 8DE1FE30 CRC32;

Query Match 74.2%, Score 49; DB 2; Length 599;
Best Local Similarity 54.5%; Pred. No. 1.44e+01;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

533 VGYLPYTAGLN 543
|::||::|::|
oy 1 VSYLTASSLD 11

RESULT	7	PRELIMINARY;	PRT;	292 AA.
ID	055748			
AC	055748:			
DT	01-JUN-1998 (TREMBLREL. 06, CREATED)			
DR	01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)			
DE	01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)			
DE	POTATIVE RIBONUCLEASE.			
OC	CHIL0 IRIDESCENT VIRUS (CIV) (INSECT IRIDESCENT VIRUS TYPE 6).			
OC	VIRUSES; DSDNA VIRUSES, NO RNA STAGE; IRIDOVIRIDAE; IRIDOVIRUS.			
RA	[1]			
RA	SEQUENCE FROM N.A.			
RA	BAHR U., TIDONA C.A., DARAI G.;			
RL	VIRUS GENES 0:0-0(1997).			
DR	EMBL; AF003534; G2738432; -			
SO	SEQUENCE 292 AA; 33664 MM; 706D89A2 CRC32;			
Query Match		72.7%;	Score 48;	DB 14; Length 292;
Best Local Similarity		45.5%;	Pred. No. 2.28e+01;	
Matches	5; Conservative		5; Mismatches	1; Indels 0; Gaps 0;
DB	39 IPLYTTPESLN 49			
	1 VSYLSTASSLD 11			
RESULT	8	PRELIMINARY;	PRT;	467 AA.
ID	043298			
AC	043298:			
DT	01-JUN-1998 (TREMBLREL. 06, CREATED)			
DT	01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)			
DT	01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)			
DE	KIAA0414.			
GN	KIAA0414 OR ZNF-X.			
OS	HOMO SAPIENS (HUMAN).			
OC	EDUAROTIA; METAOCIA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;			
OC	CATARRHINI; HOMINIDAE; HOMO.			
RA	[1]			
RA	SEQUENCE FROM N.A.			
RC	TISSUE-BRAIN:			
RA	ISHIKAWA K., NAGASE T., NAKAJIMA D., SEKI N., OHIRA M., MIYAJIMA N.,			
RA	TANAKA A., KOTANI H., NOMURA N., OHARA O.;			
RL	SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.			
RA	[2]			
RA	SEQUENCE FROM N.A.			
RA	ZHANG X., LIU C.-X., LISITSINA M.N., MUSCO S., LISITSYN N.A.;			
DR	SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.			
EMBL; AB007874; D1025766; -				

DR EMBL; AF049907; G2952308; -.
DR PROSITE; P50028; ZINC_FINGER_C2H2.
KW ZINC-FINGER; METAL-BINDING; DNA-BINDING.
SQ SEQUENCE 467 AA; 52630 MM; FA610780 CRC32;

	Query Match	Score	DB	Length
Best Local	Similarity 72.7%;	48	4	467
Matches	Conservative 7;	Pred. No. 2.28e+01;		
		Mismatches 3;	Indels 0;	Gaps 0
Dd	101 VSYLTAASTG 111			
	: :			
Oy	1 VSYLTAASTLD 11			

ID	RESULT	9	PRELIMINARY:	PRT:	832 AA.
AC	Q92831;				
DT	01-FEB-1997	(TREMBLREL, 02,	CREATED)		
DT	01-FEB-1997	(TREMBLREL, 02,	LAST SEQUENCE UPDATE)		
DT	01-NOV-1998	(TREMBLREL, 08,	LAST ANNOTATION UPDATE)		
DE	P/CAF.				
GN	P/CAF.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUAROTIA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;				
OC	EUAROTIA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;				
OC	CATARRHINI; HOMINIDAE; HOMO.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=BRAIN;				
RX	MEDLINE: 96300317.				
RA	YANG X.J., OGRZKO V.V., NISHIKAWA J., HOWARD B.H., NAKATANI Y.;				
RT	"A p300/CBP-associated factor that competes with the adenoviral				
RT	oncoprotein E1A."				
RL	NATURE 382:319-324(1996).				
DR	EMBL: U57317; G1491937; -.				
DR	PFAM: PF00439; bromodomain; 1.				
DR	PFAM: PF00583; Acetyltransf. 1.				
SO	SEQUENCE 832 AA; 92926 MW; C891243F CRC32;				

```

Query Match = 72.7%      Score 48; DB 4; Length 832;
Best Local Similarity 63.6%      Pred. No. 2.28e+01;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0

Db 394 ISYNTSSSLE 404
      ||| ||| |||:
Oy 1 VSYLSTASSDL 11

RESULT 10
ID 016568      PRELIMINARY;      PRT: 1245 AA.
AC 016568:
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE F40E12.2 PROTEIN (FRAGMENT).
GN F40E12.2.
OS CAENORHABDITIS ELEGANS.
OC EUMAROTIA; METAPOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE; 94150718.

RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DUREIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KESHAM J., KIRSTEN J., LAISTER N., LARREILLE P.,
RA LIGHTNING J., LOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PASLOS N., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRI-MEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLIDMAN P.,
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT

```

RT eiegans";
RL NATURE 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA GOELA D.;
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RL WATERSTON R.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF016559; G2315574; -.
DR PPRM: PF00023; ank; 1.
DR PPRM: PF00533; BRC1; 1.
FT NON_TER 1
SQ SEQUENCE 1245 AA; 142852 MW; A3C7849 CRC32;
Query Match 72.7%; Score 48; DB 5; Length 1245;
Best Local Similarity 45.5%; Pred. No. 2.28e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Db 290 VDFLTANSLS 300
QY 1 VSYLSTASSLD 11
RESULT 11
ID 011696 PRELIMINARY; PRT; 127 AA.
AC 011696;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE NUCLEOPROTEIN (FRAGMENT).
OS MEASLES VIRUS (SUBACUTE SCLEROSE PANENCEPHALITIS VIRUS).
OC VIRUSES: SSRNA NEGATIVE-STRAND VIRUSES: MONONEGAVIRALES;
OC PARAMYXOVIRIDAE: PARAMYXOVIRINAE: MORBILLIVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-92-E;
RX MEDLINE: 97278133.
RA YAMAGUCHI S.;
RT "Identification of three lineages of wild measles virus by nucleotide
sequence analysis of N, P, M, F, and L genes in Japan.";
RL J. MED. VIROL. 52:113-120(1997).
DR EMBL: D87487; D1020995; -.
KW NUCLEOPROTEIN.
FT NON_TER 1
SQ SEQUENCE 127 AA; 13950 MW; 42D75A2C CRC32;
Query Match 71.2%; Score 47; DB 14; Length 127;
Best Local Similarity 50.0%; Pred. No. 3.58e+01;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 67 AYLPTSTPLD 76
QY 2 SYLSTASSLD 11
RESULT 12
ID 061836 PRELIMINARY; PRT; 149 AA.
AC 061836;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MYELIN BASIC PROTEIN (FRAGMENT).
OS MUS MUSCULUS (MOUSE).
OC EURAROTIA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA: RODENTIA;
OC SCURONGNATHI: MORIDAE: MORINAE: MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-BRAIN;
RX MEDLINE: 87118269.

RA NEWMAN S., KITAMURA K., CAMPAGNONI A.T.;
RT "Identification of a cDNA coding for a fifth form of myelin basic
protein in mouse.";
RL PROC. NATL. ACAD. SCI. U.S.A. 84:886-890(1987).
DR EMBL: M15052; G199051; -.
DR PROSITE: P500569; MYELIN-MBP; 1.
KW ALTERNATIVE SPLICING; MYELIN.
FT NON_TER 1
SQ SEQUENCE 149 AA; 16226 MW; A6DC1599 CRC32;
Query Match 71.2%; Score 47; DB 11; Length 149;
Best Local Similarity 66.7%; Pred. No. 3.58e+01;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 4 YLATASTMD 12
QY 3 YLSTASSLD 11
RESULT 13
ID 001585 PRELIMINARY; PRT; 195 AA.
AC 001585;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MYELIN BASIC PROTEIN (MBP).
GN MBP.
OS MUS MUSCULUS (MOUSE).
OC EURAROTIA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA: RODENTIA;
OC SCURONGNATHI: MORIDAE: MORINAE: MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57 BL/6; TISSUE-BONE MARROW;
RX MEDLINE: 93057537.
RA GRIMA B., ZELENKA D., PESSAC B.;
RT "A novel transcript overlapping the myelin basic protein gene.";
RL J. NEUROCHEM. 59:2318-2323(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6; TISSUE-BRAIN;
RX MEDLINE: 93186801.
RA CAMPAGNONI A.T., PRIBYL T.M., CAMPAGNONI C.W., KAMPE K.,
RA ANDR-UMAJEE S., LANDRY C., HANDLEY V., NEWMAN S., GARAY B.,
RA KITAMURA K.;
RT "Structure and developmental regulation of Goli-1-mbp, a 105-kilobase
RT gene that encompasses the myelin basic protein gene and is expressed
RT in cells in the oligodendrocyte lineage in the brain.";
RL J. BIOL. CHEM. 268:4930-4938(1993).
CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO MAINTAIN PROPER STRUCTURE
CC OF MYELIN.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MYELIN.
CC EMBL: X67319; G51333; -.
DR EMBL: L07508; G193587; -.
DR MGD: MGI:96925; MBP.
KW MYELIN; STRUCTURAL PROTEIN; ACETYLATION; METHYLATION; PHOSPHORYLATION;
KW AUTOIMMUNE ENCEPHALOMYELITIS; ALTERNATIVE SPLICING.
SQ SEQUENCE 195 AA; 21004 MW; 557D83EA CRC32;
Query Match 71.2%; Score 47; DB 11; Length 195;
Best Local Similarity 66.7%; Pred. No. 3.58e+01;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 146 YLATASTMD 154
QY 3 YLSTASSLD 11
RESULT 14
ID 015339 PRELIMINARY; PRT; 197 AA.
AC 015339;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE GOLLI-MBP.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOT; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
 OC CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE; 94068468.
 RA PRIEYL T.M., CAMPAGNONI C.W., KAMPE K., KASHIMA T., HANDLEY V.W.,
 RA MCMAHON J., CAMPAGNONI A.T.;
 RT "The human myelin basic protein gene is included within a
 179-kilobase transcription unit: expression in the immune and central
 nervous systems."
 RT PROC. NATL. ACAD. SCI. U.S.A. 90:10695-10699(1993).
 RL EMBL; L18865; G435060; -
 SQ SEQUENCE 197 AA; 21522 MW; D8A8A5BF CRC32;

Query Match 71.2%; Score 47; DB 4; Length 197;
 Best Local Similarity 66.7%; Pred. No. 3.58e+01;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 3 YLSTASSLD 11
 148 YLSTASTMD 156
 ||:||||:|

RESULT 15
 ID 000510 PRELIMINARY; PRT; 239 AA.
 AC 000510;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE SRMX PROTEIN.
 GN SRMX
 OS STREPTOMYCES AMBOFACIENS.
 OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
 OC ACTINOMYCETALES; STREPTOMYCINACEAE; STREPTOMYCETACEAE; STREPTOMYCES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BES2281;
 RX MEDLINE; 92374852.
 RA GEISTLICH M., LOSICK R., TURNER J.R., RAU R.N.;
 RT "Characterization of a novel regulatory gene governing the expression
 of a polypeptide synthase gene in Streptomyces ambofaciens."
 RL MOL. MICROBIOL. 6:2019-2029(1992).
 DR EMBL; X63451; G46702; -
 SQ SEQUENCE 239 AA; 26493 MW; 45A8C51A CRC32;

Query Match 71.2%; Score 47; DB 2; Length 239;
 Best Local Similarity 36.4%; Pred. No. 3.58e+01;
 Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

DB 109 IGYLETTADLE 119
 :||| |::|:
 QY 1 VSYLSTASSLD 11

Search completed: Thu Sep 2 11:22:14 1999
 Job time : 23 secs.

THIS PAGE BLANK (USPTO)

GN MERA.
OS STAPHYLOCOCCUS AUREUS.
OG PLASMID p1258.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
CC STAPHYLOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87260937.
RA LADNAGA R.A., CHU L., MISRA T.K., SILVER S.;
RT "Nucleotide sequence and expression of the mercurial-resistance
RT operon from Staphylococcus aureus plasmid p1258."
RL PROC. NATL. ACAD. SCI. U.S.A. 84:5106-5110(1987).
CC -1- FUNCTION: RESISTANCE TO HG(2+) IN BACTERIA APPEARS TO BE GOVERNED
CC BY A SPECIALIZED SYSTEM WHICH INCLUDES MERCURY REDUCTASE. MERA
CC PROTEIN IS RESPONSIBLE FOR VOLATILIZING MERCURY AS HG(0).
CC -1- CATALYTIC ACTIVITY: HG + NADP(+) + H(+) = HG(2+) + NADPH.
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: HOMODIMER.
CC -1- THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC OXIDOREDUCTASES CLASS-I.
CC -1- SIMILARITY: CONTAINS A COPY OF THE HEAVY-METAL-ASSOCIATED (HMA)
CC DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L29436; G459907; -.
DR PIR: E29504; E29504.
DR PROSITE: PS00076; PYRIDINE_REDOX_1; 1.
DR PROSITE: PS01047; HMA; 1.
DR PFAM: PF00070; PYR_redox; 1.
DR PFAM: PF00403; HMA; 1.
DR HSSP: P11959; IEBD.
DR HSSP: P11959; IEBD.
KM MERCURY RESISTANCE; OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NADP;
KW MERCURY; REDOX-ACTIVE CENTER; METAL-BINDING; PLASMID.
FT DOMAIN 10 38 HMA
FT NP_BIND 87 117 FAD (ADP PART) (PROBABLE).
FT DISULFID 123 128 REDOX-ACTIVE.
FT NP_BIND 378 388 FAD (FLAVIN PART) (BY SIMILARITY).
FT METAL 544 544 HG(2+) (POTENTIAL).
FT METAL 545 545 HG(2+) (POTENTIAL).
NO SEQUENCE 547 AA; 58565 MW; 60DAFC29 CRC32;
DB 235 VDYLTSTALE 245
OY 1 VSYLSTASSLD 11
Query Match 78.8%; Score 52; DB 1; Length 547;
Best Local Similarity 45.5%; Pred. No. 5; 87e-01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
RESULT 3
ID YW07 MYCTU STANDARD; PRT; 132 AA.
AC Q10847;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 14.7 KD PROTEIN CY39.07C.
GN MTCY39.07C.
OS MYCOBACTERIUM TUBERCULOSIS.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
OC ACTINOMYCETALES; CORYNEBACTERIINAE; MYCOBACTERIACEAE; MYCOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA OLIVER K., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;

RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SIMILARITY: TO M.TUBERCULOSIS MTCY48.04C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 274025; E248788; -.
KM HYPOTHETICAL PROTEIN.
NO SEQUENCE 132 AA; 14731 MW; 3D577F11 CRC32;
DB 9 IAYLSTSESL 18
OY 1 VSYLSTASSLD 10
Query Match 74.2%; Score 49; DB 1; Length 132;
Best Local Similarity 60.0%; Pred. No. 3; 00e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
RESULT 4
ID YP57_CAEEL STANDARD; PRT; 468 AA.
AC Q20085;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE PUTATIVE SERINE/THREONINE-PROTEIN KINASE F35H8.7 IN CHROMOSOME II
DE (EC 2.7.1.-).
GN F35H8.7.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA BERKS M.;
RL SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP REVISIONS.
RC STRAIN-BRISTOL N2;
RA JONES S.J.M.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
CC PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 236752; E1346252; -.
DR WORKREP: F35H8.7; GE09940.
DR PROSITE: PS0107; PROTEIN_KINASE_ATP_1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM_1.
DR PFAM: PF00069; PKINase; 2.
DR HSSP: P08631; 2HCK.
KM HYPOTHETICAL PROTEIN; TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE;
KW ATP-BINDING.
FT DOMAIN 106 357 PROTEIN KINASE.
FT NP_BIND 112 121 ATP (BY SIMILARITY).
FT BINDING 136 136 ATP (BY SIMILARITY).
FT ACT_SITE 224 224 BY SIMILARITY.
NO SEQUENCE 468 AA; 54045 MW; 480A5814 CRC32;
DB 9 IAYLSTSESL 18
OY 1 VSYLSTASSLD 10
Query Match 74.2%; Score 49; DB 1; Length 468;

Best Local Similarity 60.0%; Pred. No. 3.00e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 379 ISYLTSPSL 388
:|||||:1
QY 1 VSYLSTASSL 10

RESULT 5
ID UFO1_MAIZE STANDARD; PRT: 471 AA.

AC P16166;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE FLAVONOL 3-O-GLUCOSYLTRANSFERASE (EC 2.4.1.91) (UDP-GLUCOSE FLAVONOID
3-O-GLUCOSYLTRANSFERASE) (BRONZE-1) (B2-MC2 ALLELE).
GN B21 OR UGT71A1.

OS ZEA MAYS (MAIZE).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES;
POACEAE; ZEA.

[1]
RP SEQUENCE FROM N.A.
RA FURTER D., SCHIEFELBEIN J.W., JOHNSTON F., NELSON O.E. JR.;
RT "Sequence comparisons of 3 wild-type bronze-1 alleles from Zea mays.";
RN PLANT MOL. BIOL. 11:473-481(1988).

[2]
RP SEQUENCE FROM N.A.
RA MEDLINE: 88284304.

RT RALSTON E.J., ENGLISH J.J., DOONER H.K.;
RT "Sequence of three bronze alleles of maize and correlation with the
genetic fine structure.";
RL GENETICS 119:185-197(1988).

CC -1- FUNCTION: IN THE PRESENCE OF OTHER NECESSARY COLOR FACTORS, THIS
GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
PIGMENTS.

CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + A FLAVONOL = UDP + FLAVONOL
3-O-D-GLUCOSIDE.

CC -1- PATHWAY: ONE OF THE ESSENTIAL AND TERMINAL STEPS IN THE
ANTHOCYANIN BIOSYNTHETIC PATHWAY.

CC -1- SIMILARITY: BELONGS TO THE UDP-GLUCOSYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: X13500; G1030071; .
DR EMBL: X07940; G22205; .
DR PIR: S01052; S01052.
DR PIR: S08324; S08324.
DR MAIZEDB: 13885; .
DR PROSITE: PS00375; UDEPT: 1.
DR PFAM: PF00201; UDEPT: 2.
KW TRANSFERASE; GLYCOSYLTRANSFERASE.
SQ SEQUENCE 471 AA; 48769 MW; 8AE03FD2 CRC32;

Query Match 74.2%; Score 49; DB 1; Length 471;
Best Local Similarity 80.0%; Pred. No. 3.00e+00;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 46 LSFSTASSL 55
:|||||11111
QY 1 VSYLSTASSL 10

RESULT 6
ID UFO2_MAIZE STANDARD; PRT: 471 AA.
AC P16165;
DT 01-APR-1990 (REL. 14, CREATED)

DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE FLAVONOL 3-O-GLUCOSYLTRANSFERASE (EC 2.4.1.91) (UDP-GLUCOSE FLAVONOID
3-O-GLUCOSYLTRANSFERASE) (BRONZE-1) (B2-MC2 ALLELE).
GN B21 OR UGT71A1.

OS ZEA MAYS (MAIZE).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES;
POACEAE; ZEA.

[1]
RP SEQUENCE FROM N.A.
RA FURTER D., SCHIEFELBEIN J.W., JOHNSTON F., NELSON O.E. JR.;
RT "Sequence comparisons of 3 wild-type bronze-1 alleles from Zea mays.";
RN PLANT MOL. BIOL. 11:473-481(1988).

CC -1- FUNCTION: IN THE PRESENCE OF OTHER NECESSARY COLOR FACTORS, THIS
GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
PIGMENTS.

CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + A FLAVONOL = UDP + FLAVONOL
3-O-D-GLUCOSIDE.

CC -1- PATHWAY: ONE OF THE ESSENTIAL AND TERMINAL STEPS IN THE
ANTHOCYANIN BIOSYNTHETIC PATHWAY.

CC -1- SIMILARITY: BELONGS TO THE UDP-GLUCOSYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: X13501; G295854; .
DR PIR: S08325; S08325.

DR MAIZEDB: 13885; .
DR PROSITE: PS00375; UDEPT: 1.

DR PFAM: PF00201; UDEPT: 2.
KW TRANSFERASE; GLYCOSYLTRANSFERASE.

SQ SEQUENCE 471 AA; 48621 MW; 3158C5E0 CRC32;

Query Match 74.2%; Score 49; DB 1; Length 471;
Best Local Similarity 80.0%; Pred. No. 3.00e+00;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 46 LSFSTASSL 55
:|||||11111
QY 1 VSYLSTASSL 10

RESULT 7
ID UFO3_MAIZE STANDARD; PRT: 471 AA.

AC P16167;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE FLAVONOL 3-O-GLUCOSYLTRANSFERASE (EC 2.4.1.91) (UDP-GLUCOSE FLAVONOID
3-O-GLUCOSYLTRANSFERASE) (BRONZE-1) (B2-MC2 ALLELE).
GN B21 OR UGT71A1.

OS ZEA MAYS (MAIZE).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES;
POACEAE; ZEA.

[1]
RP SEQUENCE FROM N.A.
RA FURTER D., SCHIEFELBEIN J.W., JOHNSTON F., NELSON O.E. JR.;
RT "Sequence comparisons of 3 wild-type bronze-1 alleles from Zea mays.";
RN PLANT MOL. BIOL. 11:473-481(1988).

[2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88284304.
RA RALSTON E.J., ENGLISH J.J., DOONER H.K.;
RT "Sequence of three bronze alleles of maize and correlation with the
genetic fine structure.";
RL GENETICS 119:185-197(1988).

```

CC -1- FUNCTION: IN THE PRESENCE OF OTHER NECESSARY COLOR FACTORS, THIS
CC GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
CC PIGMENTS.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + A FLAVONOL = UDP + FLAVONOL
CC 3-O-D-GLUCOSIDE.
CC -1- PATHWAY: ONE OF THE ESSENTIAL AND TERMINAL STEPS IN THE
CC ANTHOCYANIN BIOSYNTHETIC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X13503; G22506; -
DR EMBL: X07937; G22210; -
DR PIR: S01037; S01037;
DR PIR: S08326; S08326;
DR MAI2EDB: 13885; -
DR PROSITE: PS00375; UDCPT: 1.
DR PRAM: PF00201; UDCPT: 2.
KW TRANSFERASE; GLYCOSYLTRANSFERASE.
SQ SEQUENCE 471 AA; 48673 MW; 4A3C6193 CRC32;

Query Match 74.2%; Score 49; DB 1; Length 471;
Best Local Similarity 80.0%; Pred. No. 3.00e+00;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 46 LSFLSTASL 55
QY 1 VSYLSTASL 10

RESULT 8 STANDARD; PRT; 2464 AA.
AC P14873;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE MICROTUBULE-ASSOCIATED PROTEIN 1B (MAP1.2) [MAP1(X)] [CONTAINS: LIGHT
DE CHAIN LC1].
GN MAP1B OR MTAP5.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCINOGNATHI; MURIDAE; MURINAE; MUS.
[1]
SEQUENCE FROM N.A.
RC STRAIN-SWISS WEBSTER; TISSUE-BRAIN;
RX MEDLINE; 90094539.
RA NOBLE M., LEWIS S.A., COMAN N.J.;
RT "The microtubule binding domain of microtubule-associated protein
RT MAP1B contains a repeated sequence motif unrelated to that of MAP2
RT and tau."
RL J. CELL BIOL. 109:3367-3376(1989).
CC -1- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.
CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES
CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST
CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS
CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN
CC STABILIZING MICROTUBULES.
CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAP1A AND MAP1B PROTEINS.
CC -1- DOMAIN: HAS A HIGHLY BASIC REGION WITH MANY COPIES OF THE SEQUENCE
CC KKEE AND KKEI/V, REPEATED BUT NOT AT FIXED INTERVALS. THIS LATTER
CC REGION IS RESPONSIBLE FOR THE BINDING OF MAP1B TO MICROTUBULES
CC BOTH IN VITRO AND IN VIVO.
CC -1- PTM: LC1 IS COEXPRESSED WITH MAP1B. IT IS A POLYPEPTIDE GENERATED
CC FROM MAP1B BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
CC BOTH MAP1A AND MAP1B. IT INTERACTS WITH THE AMINO-TERMINAL REGION
CC OF MAP1B.

```

```

CC -1- SIMILARITY: TO NEURAXIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51396; G53000; -
DR PIR: S07549; QKMSPI.
DR MGD; MG1:97179; MTAP5.
DR PROSITE: PS00230; MAP1B_NEURAXIN; 7.
DR PRAM: PF00414; MAP1B_neuraxin; 10.
KW MICROTUBULES; REPEAT; PHOSPHORYLATION.
FT CHAIN 2 2464
FT DOMAIN 589 787
FT 1865 2068
FT REPEAT 1865 1881
FT REPEAT 1882 1898
FT REPEAT 1899 1915
FT REPEAT 1916 1932
FT REPEAT 1933 1949
FT REPEAT 1950 1966
FT REPEAT 1967 1983
FT REPEAT 1984 2000
FT REPEAT 2001 2017
FT REPEAT 2018 2034
FT REPEAT 2035 2051
FT REPEAT 2052 2068
SQ SEQUENCE 2464 AA; 270408 MW; 44FF677 CRC32;

Query Match 74.2%; Score 49; DB 1; Length 2464;
Best Local Similarity 60.0%; Pred. No. 3.00e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 463 ISYLTYSYL 472
QY 1 VSYLSTASL 10

RESULT 9 STANDARD; PRT; 2468 AA.
AC P46821;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE MICROTUBULE-ASSOCIATED PROTEIN 1B [CONTAINS: LIGHT CHAIN LC1].
GN MAP1B.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
[1]
SEQUENCE FROM N.A.
RX MEDLINE; 95104835.
RA LIEN L.T., FEENER C., FISCHBACH N., KUNKEL L.M.;
RT "Cloning of human microtubule-associated protein 1B and the
RT identification of a related gene on chromosome 15."
RL GENOMICS 22:273-280(1994).
CC -1- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.
CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES
CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST
CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS
CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN
CC STABILIZING MICROTUBULES.
CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAP1A AND MAP1B PROTEINS.
CC -1- DOMAIN: HAS A HIGHLY BASIC REGION WITH MANY COPIES OF THE SEQUENCE
CC KKEE AND KKEI/V, REPEATED BUT NOT AT FIXED INTERVALS. THIS LATTER
CC REGION IS RESPONSIBLE FOR THE BINDING OF MAP1B TO MICROTUBULES
CC BOTH IN VITRO AND IN VIVO.

```

CC -1- PTM: LC1 IS COEXPRESSED WITH MAP1B. IT IS A POLYPEPTIDE GENERATED
 CC FROM MAP1B BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
 CC BOTH MAP1A AND MAP1B.
 CC -1- SIMILARITY: TO NEURAXIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L06237; G473431; -.
 CC DR MIR: 157129; -.
 CC DR PROSITE: PS00230; MAP1B_NEURAXIN: 6.
 CC DR PFAM: PF00414; MAP1B_neuraxin: 10.
 CC KW MICROTUBULES; REPEAT; PHOSPHORYLATION.
 CC FT CHAIN ? 2468 MAP1B LC1.
 CC FT DOMAIN 589 790 LYS-RICH (HIGHLY BASIC, CONTAINS MANY
 CC KEE AND KKEI/V REPEATS).
 CC FT REPEAT 1869 2074 12 X 17 AA TANDEM REPEATS.
 CC FT REPEAT 1869 1885 1.
 CC FT REPEAT 1886 1902 2.
 CC FT REPEAT 1903 1919 3.
 CC FT REPEAT 1920 1936 4.
 CC FT REPEAT 1937 1953 5.
 CC FT REPEAT 1954 1970 6.
 CC FT REPEAT 1971 1987 7.
 CC FT REPEAT 1988 2004 8.
 CC FT REPEAT 2005 2021 9.
 CC FT REPEAT 2022 2038 10.
 CC FT REPEAT 2039 2055 11.
 CC FT REPEAT 2056 2072 12.
 CC FT REPEAT 2056 2072 12.
 CC SQ SEQUENCE 2468 AA; 270618 MW; 862C5F69 CRC32;
 CC -----
 CC Query Match 74.2%; Score 49; DB 1; Length 2468;
 CC Best Local Similarity 60.0%; Pred. No. 3.00e+00;
 CC Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 CC Db 463 ISYLTSSSL 472
 CC QY 1 VSYLTSSSL 10
 CC -----
 CC RESULT 10 STANDARD; PRT; 127 AA.
 CC ID MBP_RAT
 CC P02688;
 CC 21-JUL-1986 (REL. 01, CREATED)
 CC 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 CC 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 CC DE MYELIN BASIC PROTEIN S (MBP S).
 CC GN MBP.
 CC OS RATTUS NORVEGICUS (RAT).
 CC OC EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 CC OC RODENTIA; SCURIONATHI; MURIDAE; MURINAE; RATTUS.
 CC RN [1]
 CC RN SEQUENCE FROM N.A.
 CC RX MEDLINE: 87026249.
 CC RA SCHAICH M., BUDZINSKI R.M., STOFFEL W.;
 CC RT "Cloned proteolipid protein and myelin basic protein cDNA.
 CC RL Transcription of the two genes during myelination.";
 CC RL BIOL. CHEM. HOPPE-SEYLER 367:825-834(1986).
 CC RN [2]
 CC RN SEQUENCE FROM N.A.
 CC RX MEDLINE: 84026484.
 CC RA ROACH A., BOYLAN K., HORVATH S., PRUSINER S.B., HOOD L.E.;
 CC RT "Characterization of cloned cDNA representing rat myelin basic
 CC RL protein: absence of expression in brain of shiverer mutant mice.";
 CC RL CELL 34:799-806(1983).
 CC RN [3]
 CC RN SEQUENCE.
 CC RP MEDLINE: 75127359.
 CC -----

RA DUNKLEY P.R., CARNEGIE P.R.;
 RT "Amino acid sequence of the smaller basic protein from rat brain
 RT myelin.";
 RL BIOCHEM. J. 141:243-255(1974).
 RN [4]
 RN SEQUENCE OF 45-85.
 RX MEDLINE: 73180720.
 RA MCFARLIN D.E., BLANK S.E., KIBLER R.F., MCKNEALLY S., SHAPIRA R.;
 RT "Experimental allergic encephalomyelitis in the rat: response to
 RT encephalitogenic proteins and peptides.";
 RL SCIENCE 179:478-480(1973).
 CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO MAINTAIN PROPER STRUCTURE
 CC OF MYELIN.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MYELIN.
 CC -1- ALTERNATIVE PRODUCTS: RATS HAVE TWO MYELIN BASIC PROTEINS. THE
 CC SMALLER ONE, SHOWN HERE, IS MISSING 40 RESIDUES (FOLLOWING RESIDUE
 CC 113 OR 114) WITH RESPECT TO THE LARGER ONES FROM OTHER SPECIES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M25889; G205322; -.
 CC DR EMBL: K00512; -; NOT_ANNOTATED_CDS.
 CC DR PIR: A03142; MBRTS.
 CC DR PIR: B24351; B24351.
 CC DR PIR: A21062; A21062.
 CC DR PROSITE: PS00569; MYELIN_MBP: 1.
 CC KW MYELIN; STRUCTURAL PROTEIN; ACETYLATION; METHYLATION; PHOSPHORYLATION;
 CC KW AUTOIMMUNE ENCEPHALOMYELITIS; ALTERNATIVE SPLICING.
 CC FT INT_MET 0
 CC FT MOD_RES 1 1 ACETYLATION (MONO-:44% OR DI-:11%).
 CC FT MOD_RES 104 104 SG -> GS (IN REF. 4).
 CC FT CONFLICT 46 47
 CC FT CONFLICT 124 124 M -> I (IN REF. 2).
 CC SQ SEQUENCE 127 AA; 14080 MW; 834FEBF5 CRC32;
 CC -----
 CC Query Match 71.2%; Score 47; DB 1; Length 127;
 CC Best Local Similarity 66.7%; Pred. No. 8.55e+00;
 CC Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 CC Db 14 YLATASMD 22
 CC QY 3 YLATASMD 11
 CC -----
 CC RESULT 11 STANDARD; PRT; 167 AA.
 CC ID MBP_CAVPO
 CC P25188;
 CC 01-MAY-1992 (REL. 22, CREATED)
 CC 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 CC 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
 CC DE MYELIN BASIC PROTEIN (MBP).
 CC OS CAVIA PORCELLUS (GUINEA PIG).
 CC OC EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 CC OC RODENTIA; HYSTRICOGNATHI; CAVIIDAE; CAVIA.
 CC RN [1]
 CC RN SEQUENCE.
 CC RX MEDLINE: 84215086.
 CC RA DEIBLER G.E., MARTENSON R.E., KRUTZSCH H.C., KIES M.W.;
 CC RT "Sequence of guinea pig myelin basic protein.";
 CC RL J. NEUROCHEM. 43:100-105(1984).
 CC RN [2]
 CC RN SEQUENCE OF 45-87.
 CC RA SHAPIRA R., MCKNEALLY S.S., CHOU F., KIBLER R.F.;
 CC RT "Encephalitogenic fragment of myelin basic protein. Amino acid
 CC RL sequence of bovine, rabbit, guinea pig, monkey, and human fragments.";
 CC RL J. BIOL. CHEM. 246:4630-4640(1971).
 CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO MAINTAIN PROPER STRUCTURE

CC OF MYELIN.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MYELIN.
DR PIR: A37246; A37246.
DR PIR: C92087; C92087.
DR PROSITE: PS00569; MYELIN.MBP: 1.
KM MYELIN: STRUCTURAL PROTEIN: ACETYLATION; METHYLATION; PHOSPHORYLATION;
KM AUTOIMMUNE ENCEPHALOMYELITIS.
FT MOD_RES 1 106 ACETYLATION.
FT MOD_RES 106 106 METHYLATION (MONO- OR DI-) (BY
FT DOMAIN 114 122 SIMILARITY).
FT INDUCES EXPERIMENTAL AUTOIMMUNE
FT ENCEPHALOMYELITIS.
SQ SEQUENCE 167 AA; 18213 MW; B31793AB CRC32;

Query Match 71.2%; Score 47; DB 1; Length 167;
Best Local Similarity 66.7%; Pred. No. 8.55e+00;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 14 YLASTASMD 22
11:111111
3 YLASTASLD 11

RESULT 12
ID MBP_PANTR STANDARD: PRT: 171 AA.
AC P06906;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE MYELIN BASIC PROTEIN (MBP).
GN MBP.
OS PAN TROGLODYTES (CHIMPANZEE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
CC PRIMATES; CATARRHINI; HOMINIDAE; PAN.
RN [1]
RP PRELIMINARY SEQUENCE.
RX MEDLINE: 76009821.
RA WESTALL F.C., THOMPSON M., KALTER S.S.;
RT "The proposed sequence of the encephalitogenic protein from
RT chimpanzee brain."
RL LIFE SCI. 17:219-223(1975).
CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO MAINTAIN PROPER STRUCTURE
CC OF MYELIN.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MYELIN.
CC PIR: A03139; MBP2B.
DR PIR: PS00569; MYELIN.MBP: 1.
KM MYELIN: STRUCTURAL PROTEIN: ACETYLATION; METHYLATION; PHOSPHORYLATION;
KM AUTOIMMUNE ENCEPHALOMYELITIS.
FT MOD_RES 1 107 ACETYLATION.
FT MOD_RES 107 107 METHYLATION (BY SIMILARITY).
SQ SEQUENCE 171 AA; 18560 MW; A616427A CRC32;

Query Match 71.2%; Score 47; DB 1; Length 171;
Best Local Similarity 66.7%; Pred. No. 8.55e+00;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 14 YLASTASMD 22
11:111111
3 YLASTASLD 11

RESULT 13
ID MBP_MOUSE STANDARD: PRT: 194 AA.
AC P04370;
DT 20-MAR-1987 (REL. 04, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE MYELIN BASIC PROTEIN (MBP).
GN MBP.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
RN RODENTIA; SCIROGNATHI; MURIDAE; MORINAE; MUS.
RN [1]

RP SEQUENCE FROM N.A. (ALL FOUR FORMS).
RX MEDLINE: 86079555.
RA DE FERRA F., ENGH H., HUDSON L., KAMHOLZ J., PUCKETT C., MOLINEUX S.,
RA LAZZARINI R.A.;
RT "Alternative splicing accounts for the four forms of myelin basic
RT protein."
RL CELL 43:721-727(1985).
RN [2]
RP SEQUENCE FROM N.A. (18.5 KD FORM).
RX MEDLINE: 85254913.
RA TAKAHASHI N., ROACH A., TEPELOW D.B., PRUSINER S.B., HOOD L.;
RT "Cloning and characterization of the myelin basic protein gene from
RT mouse: One gene can encode both 14 kd and 18.5 kd MBPs by alternate
RT use of exons."
RL CELL 42:139-148(1985).
RN [3]
RP SEQUENCE FROM N.A. (17 KD FORM).
RX MEDLINE: 87118269.
RA NEWMAN S., KIRAMURA K., CAMPAGNONI A.T.;
RT "Identification of a cDNA coding for a fifth form of myelin basic
RT protein in mouse."
RL PROC. NATL. ACAD. SCI. U.S.A. 84:886-890(1987).
RN [4]
RP SEQUENCE OF 1-23 FROM N.A.
RX MEDLINE: 89252919.
RA MIURA M., TAMURA T.A., AOYAMA A., MIKOSHIBA K.;
RT "The promoter elements of the mouse myelin basic protein gene
RT function efficiently in NG108-15 neuronal/glia cells."
RL GENE 75:31-38(1989).
CC -1- FUNCTION: THIS PROTEIN MAY FUNCTION TO MAINTAIN PROPER STRUCTURE
CC OF MYELIN.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MYELIN.
CC -1- ALTERNATIVE PRODUCTS: MOUSE HAS FOUR FORMS OF MBP, 21.5KD, 18.5KD,
CC 17KD, AND 14KD. PRESENT IN RELATIVE AMOUNTS OF 1:10:3.5:35. THOSE
CC 4 FORMS ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- THE SEQUENCE SHOWN IS THAT OF THE 21.5KD FORM.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: M1533; G387414; -;
DR EMBL: M1291; G387414; JOINED.
DR EMBL: M1529; G387414; JOINED.
DR EMBL: M1530; G387414; JOINED.
DR EMBL: M1531; G387414; JOINED.
DR EMBL: M1532; G387414; JOINED.
DR EMBL: M1533; G387415; -;
DR EMBL: M1291; G387415; JOINED.
DR EMBL: M1529; G387415; JOINED.
DR EMBL: M1530; G387415; JOINED.
DR EMBL: M1531; G387415; JOINED.
DR EMBL: M1532; G387415; JOINED.
DR EMBL: M1533; G387416; -;
DR EMBL: L00398; G387416; JOINED.
DR EMBL: L00400; G387416; JOINED.
DR EMBL: L00401; G387416; JOINED.
DR EMBL: L00402; G387416; JOINED.
DR EMBL: L00403; G387416; JOINED.
DR EMBL: L00398; G387417; -;
DR EMBL: L00399; G387417; JOINED.
DR EMBL: L00400; G387417; JOINED.
DR EMBL: L00401; G387417; JOINED.
DR EMBL: L00402; G387417; JOINED.
DR EMBL: L00403; G387417; JOINED.
DR EMBL: L00398; G387418; -;
DR EMBL: L00400; G387418; JOINED.
DR EMBL: L00401; G387418; JOINED.
DR EMBL: L00402; G387418; JOINED.
DR EMBL: L00403; G387418; JOINED.

```

DR EMBL: L00404; G387419; JOINED.
DR EMBL: L00398; G387419; JOINED.
DR EMBL: L00399; G387419; JOINED.
DR EMBL: L00400; G387419; JOINED.
DR EMBL: L00401; G387419; JOINED.
DR EMBL: L00402; G387419; JOINED.
DR EMBL: L00403; G387419; JOINED.
DR EMBL: M15060; G199049; JOINED.
DR EMBL: M24410; G554196; JOINED.
DR PIR: A24772; MBMSB.
DR PIR: A26591; A26591.
DR PIR: B26591; B26591.
DR MGI: 96925; MBP.
DR PROSITE: PS00569; MYELIN_MBP.
KM MYELIN, STRUCTURAL PROTEIN, ACETYLATION; METHYLATION; PHOSPHORYLATION;
KW AUTOIMMUNE ENCEPHALOMYELITIS; ALTERNATIVE SPLICING.
FT INTMET 0
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT MOD_RES 130 130 METHYLATION (BY SIMILARITY).
FT VARSPLC 57 82 MISSING (IN 18.5 KD FORM AND 14 KD FORM).
FT VARSPLC 140 180 MISSING (IN 17 KD FORM AND 14 KD FORM).
SEQUENCE 194 AA; 21371 MW; 586D9A74 CRC32;

Query Match 71.2%; Score 47; DB 1; Length 194;
Best Local Similarity 66.7%; Pred. No. 8.55e+00;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0.

Db 12 YLTASTD 20
Oy 3 YLTASTD 11

RESULT 14
ID MBP_HUMAN STANDARD; PRT; 196 AA.
AC P02686;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
DE MYELIN BASIC PROTEIN (MBP).
DE MBP.
GN GN
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CARARRHINI; HOMINIDAE; HOMO.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE: 89302693.
RA STREICHER R., STOFFEL W.;
RT "The organization of the human myelin basic protein gene. Comparison
RT with the mouse gene.";
RL BIOL. CHEM. HOPPE-SEYLER 370:503-510(1989).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE: 87311781.
RA ROTH H.J., KRONQUIST K.E., DE ROSBO N., CRANDALL B.F.,
RA CAMPAGNONI A.T.;
RT "Evidence for the expression of four myelin basic protein variants in
RT the developing human spinal cord through cDNA cloning.";
RL J. NEUROSCI. RES. 17:321-328(1987).
RN [3]
RN SEQUENCE OF 1-58; 85-131 AND 143-196 FROM N.A.
RX MEDLINE: 86308101.
RA ROTH H.J., KRONQUIST K., PRETORIUS P.J., CRANDALL B.F.,
RA CAMPAGNONI A.T.;
RT "Isolation and characterization of a cDNA coding for a novel human
RT 17.3 kmyelin basic protein (MBP) variant.";
RL J. NEUROSCI. RES. 16:227-238(1986).
RN [4]
RN SEQUENCE OF 1-58 AND 85-196 FROM N.A.
RX MEDLINE: 86259714.
RA KAMHOULZ J., DE FERRA F., PUCKETT C., LAZZARINI R.A.;
RT "Identification of three forms of human myelin basic protein by cDNA
RT cloning.";
RL PROC. NATL. ACAD. SCI. U.S.A. 83:4962-4966(1986).

```

RP	[5]	SEQUENCE OF 1-58 AND 85-196.
RA	RA	MYELIN; 72066400.
RT	RT	CARNESIE P.R.;
RL	RL	"Amino acid sequence of the encephalitogenic basic protein from human myelin.";
RP	RP	BIOCHEM. J. 123:57-67(1971).
RA	RA	[6] SEQUENCE OF 45-58 AND 85-114, REVISIONS.
RT	RT	SHAPIRA R., MCKENALLY S.S., CHOU F., KIBLER R.F.;
RL	RL	"Encephalitogenic fragment of myelin basic protein. Amino acid sequence of bovine, rabbit, guinea pig, monkey, and human fragments.";
RP	RP	J. BIOL. CHEM. 246:4630-4640(1971).
RA	RA	[7] SEQUENCE OF 1-58 FROM N.A.
RT	RT	MYELIN; 90152679.
RL	RL	BOYLAN K.B., AYRES T.M., POPKO B., TAKAHASHI N., HOOD L.E.,
RP	RP	PRUSINER S.B.;
RA	RA	"Repetitive DNA (TGGA)n 5' to the human myelin basic protein gene: a new form of oligonucleotide repetitive sequence showing length polymorphism.";
RT	RT	GENOMICS 6:16-22(1990).
RL	RL	[8]
RP	RP	METHYLATION.
RA	RA	MYELIN; 72066401.
RT	RT	BAUDWIN G.S., CARNESIE P.R.;
RL	RL	"Isolation and partial characterization of methylated arginines from the encephalitogenic basic protein of myelin.";
RP	RP	BIOCHEM. J. 123:69-74(1971).
RA	RA	[9] FUNCTION: THIS PROTEIN MAY FUNCTION TO MAINTAIN PROPER STRUCTURE OF MYELIN.
RT	RT	-1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF MYELIN.
RL	RL	-1- ALTERNATIVE PRODUCTS: HUMANS HAVE THREE FORMS OF MBP, 21.5K, 18.5K (THE MOST ABUNDANT), AND 17.2K. THE THREE FORMS ARE PRODUCED BY ALTERNATIVE SPLICING.
RP	RP	-----
RA	RA	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
RT	RT	-----
RL	RL	EMBL; M31577; G307160; -
RP	RP	EMBL; M30516; G307161; -
RA	RA	EMBL; M30515; G307162; -
RT	RT	EMBL; M30047; G307159; -
RL	RL	EMBL; X17286; E221974; -
RP	RP	EMBL; X17287; E221974; JOINED.
RA	RA	EMBL; X17290; E221974; JOINED.
RT	RT	EMBL; X17288; E221974; JOINED.
RL	RL	EMBL; X17369; E221974; JOINED.
RP	RP	EMBL; X17289; E221974; JOINED.
RA	RA	EMBL; M63599; G187403; -
RT	RT	PIR; A24153; MBHUB.
RL	RL	PIR; S10482; S10482.
RP	RP	MIM; 159430; -
RA	RA	PROSITE; PS00569; MYELIN_MBP; 1.
RT	RT	MYELIN STRUCTURAL PROTEIN; ACETYLATION; METHYLATION; PHOSPHORYLATION; AUTOIMMUNE ENCEPHALOMYELITIS; ALTERNATIVE SPLICING.
RL	RL	INIT_MBT 0
RP	RP	MOD_RES 1 1
RA	RA	MOD_RES 133 133
RT	RT	DOMAIN 45 114
RL	RL	DOMAIN 138 148
RP	RP	DOMAIN 59 84
RA	RA	VARSPLIC 132 142
RT	RT	VARSPLIC 196 AA; 21362 MW; 0A8864EE CRC32;
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN 17.2 KD FORM).
RP	RP	MISSING (IN 17.2 KD FORM).
RA	RA	MISSING (IN 17.2 KD FORM).
RT	RT	MISSING (IN 17.2 KD FORM).
RL	RL	MISSING (IN

Query Match 71.2%; Score 47; DB 1; Length 196;
 Best Local Similarity 66.7%; Pred. No. 8.55e+00;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

DB 14 YLATASTD 22
 ||:||||:|
 QY 3 YLSTASSLD 11

RESULT 15
 ID RNPH MYCTU STANDARD; PRT; 259 AA.
 AC Q10628; 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE PROBABLE RIBONUCLEASE PH (EC 2.7.7.56) (RNASE PH) (TRNA
 DE NUCLEOTIDYLTRANSFERASE)
 GN RPH OR MTCY130.25 OR MTCY02B10.04.
 NC MYCOBACTERIUM TUBERCULOSIS.
 NC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
 NC ACTINOMYCETALES; CORYNEBACTERIINEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDJ DATA BANKS.
 CC -!- FUNCTION: RNASE PH IS A PHOSPHORYLYTIC EXORIBONUCLEASE THAT
 CC REMOVES NUCLEOTIDE RESIDUES FOLLOWING THE -CCA TERMINUS OF TRNA
 CC AND ADDS NUCLEOTIDES TO THE ENDS OF RNA MOLECULES BY USING
 CC NUCLEOSIDE DIPHOSPHATES AS SUBSTRATES (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: TRNA(N+1) + ORTHOPHOSPHATE = TRNA(N) +
 CC A NUCLEOSIDE DIPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE RNASE PH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z73902; E245034; -;
 DR EMBL; Z75555; E250346; -;
 DR PROSITE; PS01277; RIBONUCLEASE_PH; 1.
 DR PFMW; PF01138; RNase_PH; 1.
 DR TRANSFERASE; NUCLEOTIDYLTRANSFERASE; TRNA PROCESSING.
 DR SEQUENCE 259 AA; 27351 MW; 89F9FE6 CRC32;

Query Match 71.2%; Score 47; DB 1; Length 259;
 Best Local Similarity 60.0%; Pred. No. 8.55e+00;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 143 VYLSAAGKL 152
 ||:||||:|
 QY 1 VYLSSTASSL 10

Search completed: Thu Sep 2 11:21:33 1999
 Job time : 8 secs.

 W E S E R E (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSearch_pp protein - protein database search, using Smith-Waterman algorithm
 on: Thu Sep 2 11:20:54 1999; Maspar time 3.25 Seconds
 147.929 Million cell updates/sec
 Molecular output not generated.

Title: >US-08-599-226-4
 Description: (1-12) from US08599226.pep
 Perfect Score: 66
 Sequence: 1 VSYLSTASSLDX 12

Scoring table: PAM 150
 Gap 15

Searched: 122810 segs, 40068593 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: p1r60
 1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 23.694; Variance 27.579; scale 0.859

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	52	78.8	547	2	E29504	mercury(II) reductase
2	51	77.3	236	2	S60941	probable membrane pro
3	50	75.8	293	2	D70108	conserved hypotheticala
4	49	74.2	132	2	H70759	hypothetical protein
5	49	74.2	300	2	A71645	protein p34 (p34) RP8
6	49	74.2	471	2	S08325	flavonol 3-O-glucosyl
7	49	74.2	471	2	S01052	flavonol 3-O-glucosyl
8	49	74.2	471	2	S01037	flavonol 3-O-glucosyl
9	49	74.2	599	2	G71481	flavonol 3-O-glucosyl
10	49	74.2	2364	2	A56577	probable phosphoenolp
11	49	74.2	2464	2	ORMSP1	microtubule-associate
12	48	72.7	292	2	T03085	microtubule-associate
13	48	72.7	832	2	S71788	ribonuclease homolog
14	47	71.2	128	1	MBRTS	P/GAF protein - human
15	47	71.2	167	1	A37246	myelin basic protein
16	47	71.2	171	1	MBCBZ	myelin basic protein
17	47	71.2	239	2	S35204	myelin basic protein
18	47	71.2	239	2	S35204	myelin basic protein
19	47	71.2	259	2	G70739	stmx protein - Strept
20	47	71.2	259	2	G72788	probable rpha protein
21	47	71.2	328	1	MBMSB	tRNA nucleotidyltrans
22	47	71.2	1107	2	E55226	goli1-myein basic pr
23	46	69.7	133	2	C71485	hypothetical 123.8 kd
						hypothetical protein

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
24	46	69.7	474	2	S30168	mercury(II) reductase
25	46	69.7	485	2	B37855	glucose-6-phosphate 1
26	46	69.7	500	2	B64680	anthranilate synthase
27	46	69.7	500	2	G71836	anthranilate synthase
28	46	69.7	554	2	J00188	membrane protein 4.1
29	46	69.7	564	2	S57124	CIP synthase (EC 6.3.
30	46	69.7	687	2	A41905	ferric vibriobactin r
31	46	69.7	761	2	C64813	YbhJ protein - Escher
32	45	68.2	87	3	D42603	ORF 3' to arbh - Erwi
33	45	68.2	261	2	A71192	probable glucose 1-de
34	45	68.2	333	2	S74406	hypothetical protein
35	45	68.2	416	2	H64984	hypothetical 43.4 kd
36	45	68.2	471	2	C71439	hypothetical protein
37	45	68.2	505	2	S76229	hypothetical protein
38	45	68.2	565	2	G33157	hemagglutinin precurs
39	45	68.2	608	2	T02684	DNA-binding protein C
40	45	68.2	619	2	A36519	CDC16 protein - human
41	45	68.2	631	2	E32227	mercury(II) reductase
42	45	68.2	633	2	S49788	probable membrane pro
43	45	68.2	650	2	F70974	probable acral protei
44	45	68.2	1369	2	JC4860	protein-tyrosine kina
45	45	68.2	1545	2	S48824	probable membrane pro

ALIGNMENTS

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	66	100.0	12	1	E29504	mercury(II) reductase (EC 1.16.1.1) - Staphylococcus aureus
2	66	100.0	12	1	mercury(II) reductase	
3	66	100.0	12	1	plasmid p1258	
4	66	100.0	12	1	Hq(II) reductase; mercuric reductase	
5	66	100.0	12	1	#formal name Staphylococcus aureus	
6	66	100.0	12	1	21-May-1988 #sequence_revision 21-May-1988 #text_change	
7	66	100.0	12	1	05-Dec-1998	
8	66	100.0	12	1	E29504	
9	66	100.0	12	1	A94170	Laddaga, R.A.; Chu, L.; Misra, T.K.; Silver, S.
10	66	100.0	12	1	Proc. Natl. Acad. Sci. U.S.A. (1987) 84:5106-5110	
11	66	100.0	12	1	Nucleotide sequence and expression of the mercurial	
12	66	100.0	12	1	resistance operon from Staphylococcus aureus plasmid p1258.	
13	66	100.0	12	1	#cross-references M01D:87260937	
14	66	100.0	12	1	#accession E29504	
15	66	100.0	12	1	#molecule_type DNA	
16	66	100.0	12	1	##residues 1-547 ##label LAD	
17	66	100.0	12	1	GENETICS	
18	66	100.0	12	1	#gene	merA
19	66	100.0	12	1	#genome	plasmid
20	66	100.0	12	1	CLASSIFICATION	#superfamily Bacillus mercury(II) reductase; dihydroliipoamide
21	66	100.0	12	1	KEYWORDS	dehydrogenase homology; heavy-metal-associated homology
22	66	100.0	12	1	FEATURE	disulfide bond; FAD; oxidoreductase
23	66	100.0	12	1	10-38	
24	66	100.0	12	1	89-528	
25	66	100.0	12	1	SUMMARY	#domain heavy-metal-associated homology #label HMA\
26	66	100.0	12	1	Query Match	#domain dihydroliipoamide dehydrogenase homology #label
27	66	100.0	12	1	Best Local Similarity 45.5%; Pred. No. 2.07e+00;	
28	66	100.0	12	1	Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;	
29	66	100.0	12	1	DB 235 VDYLFSTSALE 245	
30	66	100.0	12	1	QY 1 VSYLSTASSLD 11	
31	66	100.0	12	1	RESULT 2	
32	66	100.0	12	1	ENTRY	S60941 #type complete
33	66	100.0	12	1	TITLE	probable membrane protein YOR214c - yeast (Saccharomyces
34	66	100.0	12	1	ALTERNATE_NAMES	cerevisiae)
35	66	100.0	12	1	ORGANISM	hypothetical protein O5022; hypothetical protein YOR50-4
36	66	100.0	12	1	DATE	#formal_name Saccharomyces cerevisiae
37	66	100.0	12	1		15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change
38	66	100.0	12	1		05-Jun-1998

ACCESSIONS S60941; S67107; S71716
REFERENCE S60938
#authors Galisson, F.; Dujon, B.
#submission Submitted to the EMBL Data Library, October 1995
#description Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV of the yeast *Saccharomyces cerevisiae*.
#accession S60941
##molecule_type DNA
##residues 1-236 ##label GAL
##cross-references EMBL:X92441; NID:g1050762; PID:g1050766
REFERENCE S67104
#authors Boyer, J.; Fairhead, C.; Gallion, L.; Galisson, F.; Michaux, G.; Inierity, A.; Dujon, B.
#submission Submitted to the Protein Sequence Database, July 1996
#accession S67107
##molecule_type DNA
##residues 1-236 ##label BOY
##cross-references EMBL:Z75122; NID:g1420498; PID:e252081; PID:g1420499; MIPS:FOR214C
##experimental_source strain S288C
REFERENCE S71713
#authors Galisson, F.; Dujon, B.
#journal Yeast (1996) 12:877-885
#title Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV of the yeast *Saccharomyces cerevisiae*.
#cross-references MUID:96437977
#accession S71716
##status nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-236 ##label GAW
##cross-references EMBL:X92441; NID:g1050762; PID:g1050766
#note The nucleotide sequence was submitted to the EMBL Data Library, October 1995
GENETICS
#map_position 15R
KEYWORDS transmembrane protein
FEATURE
4-20
217-233
SUMMARY
#domain transmembrane #status predicted #label TM1\
#domain transmembrane #status predicted #label TM2
#length 236 #molecular-weight 26156 #checksum 8999
Query Match 77.3%; Score 51; DB 2; Length 236;
Best Local Similarity 60.0%; Pred. No. 3.40e+00;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Db 175 ISYLSSTSL 184
:||||:|
1 VSYLSTASL 10
RESULT 3
ENTRY D70108 #type complete
TITLE conserved hypothetical protein BB0068 - Lyme disease
ORGANISM #formal_name *Borrelia burgdorferi* #common_name Lyme disease
#formal_name *Spirochaete*
#formal_name *Spirochaete*
DATE 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 05-Jun-1998
ACCESSIONS D70108
REFERENCE A70100
#authors Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lachigra, R.; White, O.; Ketchum, K.A.; Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb, J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.; Kervage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman, J.; Uterback, T.; Wathey, L.; McDonald, L.; Artiach, P.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.
#journal Nature (1997) 390:580-586
#title Genomic sequence of a Lyme disease *Spirochaete*, *Borrelia burgdorferi*.
#cross-references MUID:96065943

#accession D70108
#status preliminary; nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-293 ##label KLE
##cross-references GB:AE001120; GB:AE000783; NID:g2687951; PID:g2687956; TIGR:BB0068
##experimental_source strain B31
SUMMARY #length 293 #molecular-weight 33278 #checksum 5223
Query Match 75.8%; Score 50; DB 2; Length 293;
Best Local Similarity 60.0%; Pred. No. 5.54e+00;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Db 198 AYLSPTNSLE 207
:||||:|
2 SYLSTASLD 11
RESULT 4
ENTRY H70759 #type complete
TITLE hypothetical protein Rv2010 - *Mycobacterium tuberculosis* (strain H37Rv)
ORGANISM #formal_name *Mycobacterium tuberculosis*
DATE 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Jul-1998
ACCESSIONS H70759
REFERENCE A70500
#authors Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gao, S.; Barry, C.E.; Tekala, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seege, K.; Skelton, S.; Squares, S.; Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
#journal Nature (1998) 393:537-544
#title Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence.
#cross-references MUID:96295987
#accession H70759
##status preliminary; nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-132 ##label COL
##cross-references GB:Z74025; GB:AL123456; NID:g3261586; PID:e248786; PID:g1403443
##experimental_source strain H37Rv
GENETICS
#gene Rv2010
SUMMARY #length 132 #molecular-weight 14731 #checksum 1326
Query Match 74.2%; Score 49; DB 2; Length 132;
Best Local Similarity 60.0%; Pred. No. 8.98e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 9 IAYLSTESL 18
:||||:|
1 VSYLSTASL 10
RESULT 5
ENTRY A71645 #type complete
TITLE protein p34 (p34) RP832 - *Rickettsia prowazekii*
ORGANISM #formal_name *Rickettsia prowazekii*
DATE 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 21-Nov-1998
ACCESSIONS A71645
REFERENCE A71630
#authors Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark, U.C.M.; Podowski, R.M.;

Neslund, A.K.; Eriksson, A.S.; Winkler, H.H.; Kurland, C.G.
Nature (1998) 396:133-140
The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
#journal
#title
#accession A71645
#status preliminary; nucleic acid sequence not shown;
translation not shown
##molecule_type DNA
##residues 1-300 ##label AND
##cross-references GB:AJ235273; GB:AJ235269; NID:g3861237; PID:e1343102;
PID:g3861358
##experimental_source strain Madrid E

GENETICS
#gene
SUMMARY p34; RP832
#length 300 #molecular-weight 34582 #checksum 8509

Query Match
Best Local Similarity 60.0%; Pred. No. 8.98e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

12 VSX1STAL 21
|||||:::
QY 1 VSX1STASL 10

RESULT 6
ENTRY S08325 #type complete
TITLE Flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (allele B2-MC2) - maize
ALTERNATE_NAMES UDPglucose flavonoid glucosyl-transferase
ORGANISM #formal_name Zea mays #common_name maize
DATE 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 08-Sep-1997

ACCESSIONS S08325
REFERENCE S08324
#authors Furek, D.; Schiefelbein, J.W.; Johnston, F.; Nelson Jr., O.E.
#journal Plant Mol. Biol. (1988) 11:473-481
#title Sequence comparisons of three wild-type Bronze-1 alleles from Zea mays.
#accession S08325
#status translation not shown
##molecule_type DNA
##residues 1-471 ##label FUR
##cross-references EMBL:X13501; NID:g22361; PID:g295854

GENETICS
#gene B21
#introns 175/1
CLASSIFICATION #superfamily flavonol O3-glucosyltransferase
KEYWORDS glycosyltransferase; hexosyltransferase
SUMMARY #length 471 #molecular-weight 48621 #checksum 7439

Query Match
Best Local Similarity 80.0%; Pred. No. 8.98e+00;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 46 LSF1STASL 55
:|||||
QY 1 VSX1STASL 10

RESULT 7
ENTRY S01052 #type complete
TITLE Flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (allele B2-MC) - maize
ALTERNATE_NAMES UDPglucose flavonoid glucosyltransferase
ORGANISM #formal_name Zea mays #common_name maize
DATE 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 08-Sep-1997

ACCESSIONS S01052; S08324
REFERENCE S01037
#authors Ralston, E.J.; English, J.J.; Dooner, H.K.

#journal Genetics (1988) 119:185-197
#title Sequence of three bronze alleles of maize and correlation with the genetic fine structure.
#cross-references MUID:88284304
#accession S01052
#status translation not shown
##molecule_type DNA
##residues 1-471 ##label RAL
##cross-references EMBL:X07940; NID:g22204; PID:g22205
S08324
#authors Furek, D.; Schiefelbein, J.W.; Johnston, F.; Nelson Jr., O.E.
#journal Plant Mol. Biol. (1988) 11:473-481
#title Sequence comparisons of three wild-type Bronze-1 alleles from Zea mays.
#accession S08324
#status translation not shown
##molecule_type DNA
##residues 1-471 ##label FUR
##cross-references EMBL:X13500; NID:g22364; PID:g1030071

GENETICS
#gene B21
#map_position 9
#introns 175/1
CLASSIFICATION #superfamily flavonol O3-glucosyltransferase
KEYWORDS glycosyltransferase; hexosyltransferase
SUMMARY #length 471 #molecular-weight 48769 #checksum 6660

Query Match
Best Local Similarity 74.2%; Score 49; DB 2; Length 471;
Pred. No. 8.98e+00;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 46 LSF1STASL 55
:|||||
QY 1 VSX1STASL 10

RESULT 8
ENTRY S01037 #type complete
TITLE Flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (allele B2-W22) - maize
ALTERNATE_NAMES UDPglucose flavonoid glucosyltransferase
ORGANISM #formal_name Zea mays #common_name maize
DATE 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 08-Sep-1997

ACCESSIONS S01037; S08326
REFERENCE S01037
#authors Ralston, E.J.; English, J.J.; Dooner, H.K.
#journal Genetics (1988) 119:185-197
#title Sequence of three bronze alleles of maize and correlation with the genetic fine structure.
#cross-references MUID:88284304
#accession S01037
#status translation not shown
##molecule_type DNA
##residues 1-471 ##label RAL
##cross-references EMBL:X07937; NID:g22209; PID:g22210
S08324
#authors Furek, D.; Schiefelbein, J.W.; Johnston, F.; Nelson Jr., O.E.
#journal Plant Mol. Biol. (1988) 11:473-481
#title Sequence comparisons of three wild-type Bronze-1 alleles from Zea mays.
#accession S08326
#status translation not shown
##molecule_type RNA
##residues 1-471 ##label FUR
##cross-references EMBL:X13502; NID:g22505; PID:g22506

GENETICS
#gene B21
#map_position 9
#introns 175/1
CLASSIFICATION #superfamily flavonol O3-glucosyltransferase

KEYWORDS anthocyanin biosynthesis; glycosyltransferase;
hexosyltransferase
SUMMARY #length 471 #molecular-weight 48673 #checksum 7590

Query Match 74.2%; Score 49; DB 2; Length 471;
Best Local Similarity 80.0%; Pred. No. 8.98e+00;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 46 LSFSTASL 55
QY 1 VSYLSTASL 10

RESULT 9
ENTRY G71481 #type complete
TITLE probable phosphoenolpyruvate carboxykinase - Chlamydia
trachomatis (serotype D, strain UW3/CX)
ORGANISM #formal_name Chlamydia trachomatis
DATE 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change
21-Nov-1998

ACCESSIONS
REFERENCE
#authors Stephens, R.S.; Kaiman, S.; Lammel, C.J.; Fan, J.; Marathe,
R.; Aravind, L.; Mitchell, W.P.; Olinger, L.; Tatusov,
R.L.; Zhao, Q.; Koplin, E.V.; Davis, R.W.
#journal Science (1998) 282:734-739
#title Genome sequence of an obligate intracellular pathogen of
humans: Chlamydia trachomatis.
#cross-references MUID:9900809
#accession G71481
#status preliminary
#molecule_type DNA
#residues 1-599 ##label ARN
#cross-references GB:AE001341; GB:AE001273; NID:g3329156; PID:g3329165
##experimental_source serotype D, strain UW-3/CX

GENETICS
#gene pckA
CLASSIFICATION #superfamily phosphoenolpyruvate carboxykinase (GTP)
SUMMARY #length 599 #molecular-weight 66244 #checksum 5490

Query Match 74.2%; Score 49; DB 2; Length 599;
Best Local Similarity 54.5%; Pred. No. 8.98e+00;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 533 VGYLPTAEGLN 543
QY 1 VSYLSTASL 11

SULT 10
ENTRY A56577 #type fragment
TITLE microtubule-associated protein MAP 1B - rat (fragment)
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change
16-Feb-1997

ACCESSIONS
REFERENCE
#authors Zsuner, W.; Kralz, J.; Staunton, J.; Felck, P.; Wiche, G.
#journal Eur. J. Cell Biol. (1992) 57:66-74
#title Identification of two distinct microtubule binding domains on
recombinant rat MAP 1B.
#cross-references MUID:92347374
#accession A56577
#status preliminary
#molecule_type mRNA
#residues 1-2364 #label ZAV
#cross-references GB:X60550
#experimental_source Brain
#note nucleotide sequence not given; conceptual translation
not complete
CLASSIFICATION #superfamily microtubule-associated protein MAP1B
SUMMARY #length 2364 #checksum 9159

Query Match 74.2%; Score 49; DB 2; Length 2364;
Best Local Similarity 60.0%; Pred. No. 8.98e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 367 ISYLSVSSL 376
QY 1 VSYLSTASL 10

RESULT 11
ENTRY ORMSP1 #type complete
TITLE microtubule-associated protein MAP1B - mouse
ALTERNATE_NAMES microtubule-associated protein MAP1(X);
microtubule-associated protein MAP1.2;
microtubule-associated protein MAP5
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change
05-Sep-1997

ACCESSIONS
REFERENCE
#authors Noble, M.; Lewis, S.A.; Cowan, N.J.
#journal J. Cell Biol. (1989) 109:3367-3376
#title The microtubule binding domain of microtubule-associated
protein MAP1B contains a repeated sequence motif unrelated
to that of MAP2 and tau.
#cross-references MUID:90094539
#accession S07549
#molecule_type mRNA
#residues 1-2464 ##label NOB
CLASSIFICATION #superfamily microtubule-associated protein MAP1B
KEYWORDS microtubule binding; phosphoprotein; tandem repeat
FEATURE
589-786 #domain microtubule binding #status experimental #label
MTB\

589-592,639-642,
649-652,655-658,
660-663,668-671,
674-677,679-682,
683-686,687-690,
691-694,695-698,
699-702,708-711,
712-715,716-719,
720-723,727-730,
738-761,764-767,
783-786
1861-2064
91,116,351,888,
1124,1153,1168,
1208,1662,1877,
1918,2003,2030,
2054,2083
147,969,1336,1562,
1563,1702,1708,
1990,2057,2063,
2419
1953
SUMMARY #length 2464 #molecular-weight 270408 #checksum 2525

Query Match 74.2%; Score 49; DB 1; Length 2464;
Best Local Similarity 60.0%; Pred. No. 8.98e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 463 ISYLSVSSL 472
QY 1 VSYLSTASL 10

RESULT 12
ENTRY T03085 #type complete

TITLE ribonuclease homolog 068R - Chilio iridescent virus
ORGANISM #formal_name Chilio iridescent virus
DATE 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999

ACCESSIONS
REFERENCE
#authors Baht, U.; Tidon, C.A.; Darai, G.
#journal Virus Genes (1997) 15:235-245
#title The DNA sequence of Chilio iridescent virus between the genome coordinates 0.101 and 0.391: similarities in coding strategy between insect and vertebrate iridoviruses.
#accession T03085
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-292 ##label BAH
#cross-references EMBL:AF003534; NID:g2738385; PID:g2738432
SUMMARY #length 292 #molecular-weight 33664 #checksum 532

Query Match 72.7%; Score 48; DB 2; Length 292;
Best Local Similarity 45.5%; Pred. No. 1.45e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 39 IPYITPESLN 49
1 VSYLSTASLD 11

RESULT 13
ENTRY S71788 #type complete
TITLE P/CAF protein - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 02-Jul-1998

ACCESSIONS
REFERENCE
#authors Yang, X.J.; Ogrzyko, V.V.; Nishikawa, J.; Howard, B.H.; Nakatani, Y.
#journal Nature (1996) 382:319-324
#title A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A.
#cross-references M01D:96300317
#accession S71788
#status Preliminary; nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-832 ##label YAN
#cross-references EMBL:U57317; NID:g1491936; PID:g1491937
CLASSIFICATION #superfamily bromodomain homology
FEATURE 748-803
#domain bromodomain homology #label BRO1
#length 832 #molecular-weight 92926 #checksum 8709

Query Match 72.7%; Score 48; DB 2; Length 832;
Best Local Similarity 63.6%; Pred. No. 1.45e+01;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 394 ISYNTSSSLE 404
1 VSYLSTASLD 11

RESULT 14
ENTRY MBRTS #type complete
TITLE myelin basic protein S-rat
ALTERNATE_NAMES small myelin basic protein
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 24-Apr-1984 #sequence_revision 08-Feb-1996 #text_change 05-Sep-1997
ACCESSIONS
REFERENCE
#authors Schach, M.; Budzinski, R.M.; Stoffel, W. Biot. Chem. Hoppe-Seyler (1986) 367:825-834
#journal Cloned proteolipid protein and myelin basic protein cDNA. Transcription of the two genes during myelination.

#cross-references M01D:87026249
#accession B24351
#molecule_type mRNA
#residues 1-128 ##label SCH
#cross-references EMBL:M25889; NID:g205321; PID:g205322
REFERENCE
#authors A90275
#journal Dunkley, P.R.; Carnegie, P.R. Blochem. J. (1974) 141:243-255
#title Amino acid sequence of the smaller basic protein from rat brain myelin.
#cross-references M01D:75127359
#accession A90275
#molecule_type protein
#residues 2-128 ##label DUN
#note at position 105, arginine, monomethylarginine, and dimethylarginine occur in the ratio 4:4:1
#note rats have two myelin basic proteins; the smaller one, shown above, is missing 40 residues (following residue 114 or 115) with respect to the larger ones from other species

REFERENCE
#authors A94243
#journal McFarlin, D.E.; Blank, S.E.; Kibler, R.F.; McKneally, S.; Shapira, R. Science (1973) 179:478-480
#title Experimental allergic encephalomyelitis in the rat: response to encephalitogenic proteins and peptides.
#cross-references M01D:73180720
#accession A94243
#molecule_type protein
#residues 46-86 ##label MCF
#note the sequence reported for this encephalitogenic peptide differs from that shown by a transposition of residues 47 and 48; two other differences are printing errors

REFERENCE
#authors A21062
#journal Roach, A.; Boylan, K.; Horvath, S.; Prusiner, S.B.; Hood, L.E. Cell (1983) 34:799-806
#title Characterization of cloned cDNA representing rat myelin basic protein: absence of expression in brain of shiverer mutant mice.
#cross-references M01D:84026484
#accession A21062
#molecule_type mRNA
#residues 1-124, '1', 126-128 ##label ROA
#experimental_source strain Sprague-Dawley
#classification #superfamily myelin basic protein
#keywords alternative splicing; blocked amino end; experimental autoimmune encephalomyelitis; methylated amino acid; myelin

FEATURE 2-128
#product myelin basic protein S #status experimental
#label MAT\

2 #modified_site blocked amino end (Ala) (in mature form) (probably acetylated) #status experimental
105 #modified_site omega-N-methylarginine or omega-N, omega-N'-dimethylarginine (Arg) (partial) #status experimental

SUMMARY #length 128 #molecular-weight 14211 #checksum 2812

Query Match 71.2%; Score 47; DB 1; Length 128;
Best Local Similarity 66.7%; Pred. No. 2.31e+01;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 15 YLATASTMD 23
3 YLSTASLD 11

RESULT 15
ENTRY A37246 #type complete
TITLE myelin basic protein - guinea pig
ALTERNATE_NAMES myelin AI protein
ORGANISM #formal_name Cavia porcellus #common_name guinea pig
DATE 31-Jul-1991 #sequence_revision 31-Jul-1991 #text_change

07-Oct-1994
ACCESSIONS A37246; C92087; A03140
REFERENCE A37246
#authors Deibler, G.E.; Martenson, R.E.; Krutzsch, H.C.; Kies, M.W.
#journal J. Neurochem. (1984) 43:100-105
#title Sequence of guinea pig myelin basic protein.
#cross-references MUID:84215086
#accession A37246 preliminary
#status
#molecule-type protein
#residues 1-167 #label DE1
REFERENCE A92087
#authors Shapiro, R.; McKeally, S.S.; Chou, F.; Kibler, R.F.
#journal J. Biol. Chem. (1971) 246:4630-4640
#title Encephalitogenic fragment of myelin basic protein. Amino acid
sequence of bovine, rabbit, guinea pig, monkey, and human
fragments.
#accession C92087
#molecule-type protein
#residues 45-87 #label SMA
#classification #superfamily myelin basic protein
#keywords myelin
SUMMARY #length 167 #molecular-weight 18213 #checksum 1628
Query Match 71.2%; Score 47; DB 2; Length 167;
Best Local Similarity 66.7%; Pred. No. 2.31e+01;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 14 YLATASTD 22
||:||||:|
OY 3 YLSTASLD 11

Search completed: Thu Sep 2 11:21:08 1999
Job time : 14 secs.

Sequence 426 AA; 47234 MW; 1032622 CN; SQ

Query Match 72.7%; Score 48; DB 3; Length 426;
Best Local Similarity 70.0%; Pred. No. 5.13e+01;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 215 SYLSPSPPLD 224
||| :|:|
Qy 2 SYLSTASSLD 11

RESULT 2 STANDARD; PRT; 426 AA.

XX US-08-336-583-2

Sequence 2, Application US/08336583

Patent No. 5629415

GENERAL INFORMATION:

APPLICANT: HOLDIS, GREGORY F.

APPLICANT: PATEL, MAYUR D.

TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN E

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: CHRISTINE E. CARTY

STREET: 126 E. LINCOLN AVENUE

CITY: RAHWAY

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/336,583

FILING DATE: 09-NOV-1994

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: CARTY, CHRISTINE E.

REGISTRATION NUMBER: 36,099

REFERENCE/DOCKET NUMBER: 19211

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 594-6734

TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 426 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 426 AA; 47234 MW; 1032622 CN;

Query Match 72.7%; Score 48; DB 1; Length 426;
Best Local Similarity 70.0%; Pred. No. 5.13e+01;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 215 SYLSPSPPLD 224
||| :|:|

Qy 2 SYLSTASSLD 11

RESULT 3 STANDARD; PRT; 20 AA.

XX US-08-468-540B-2

AC xxxxxx
XX
DT
XX

DE Sequence 2, Application US/08468540B

XX Patent No. 5858980

CC Sequence 2, Application US/08468540B

CC GENERAL INFORMATION:

CC APPLICANT: Welner, Howard

CC APPLICANT: Haller, David

CC APPLICANT: Miller, Ariel

CC APPLICANT: Al-Sabbagh, Ahmad

CC TITLE OF INVENTION: SUPPRESSION OF T-CELL PROLIFERATION

CC TITLE OF INVENTION: USING PEPTIDE FRAGMENTS OF MYELIN BASIC PROTEIN

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Darby & Darby P.C.

STREET: 805 Third Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,540B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Jacobs, Seth H

REGISTRATION NUMBER: 32,140

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-527-7700

TELEFAX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 5858980e

SEQUENCE 20 AA; 2160 MW; 2047 CN;

Query Match 71.2%; Score 47; DB 2; Length 20;
Best Local Similarity 66.7%; Pred. No. 6.62e+01;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 4 YIATASTMD 12
||:||||:|

Qy 3 YIATASSLD 11

RESULT 4 STANDARD; PRT; 21 AA.

XX US-08-787-547-33

Sequence 33, Application US/08787547

Patent No. 5783567

GENERAL INFORMATION:

APPLICANT: Hedley, Mary Lynne

APPLICANT: Curley, Joanne M.

APPLICANT: Langer, Robert S.

TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY

CC TITLE OF INVENTION: OF NUCLEIC ACID
CC NUMBER OF SEQUENCES: 107
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson, P.C.
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: MA
CC COUNTRY: US
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: Windows95
CC SOFTWARE: FASTSEQ for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/787,547
CC FILING DATE: 22-JAN-1997
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fraser, Janis K.
CC REGISTRATION NUMBER: 34,819
CC REFERENCE/DOCKET NUMBER: 08191/003001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-542-5070
CC TELEFAX: 617-542-8906
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 33:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 21 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 21 AA; 2312 MW; 2012 CN;
SQ
Query Match 71.2%; Score 47; DB 2; Length 21;
Best Local Similarity 66.7%; Pred. No. 6.62e+01;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 8 YLATASTMD 16
QY 3 YLISTASSLD 11
RESULT 5
5194425-4 STANDARD; PRT: 182 AA.
xxxxxx
XX 01-JAN-1900
XX Patent No. 5194425.
XX Patent No. 5194425.
XX APPLICANT: SHARMA, SOMESH D.; LERCH, L. BERNARD; CLARK, BRIAN R.
XX TITLE OF INVENTION: MHC-MEDIATED TOXIC CONJUGATES USEFUL IN AMELIORATING AUTOIMMUNITY
XX NUMBER OF SEQUENCES: 9
XX CURRENT APPLICATION DATA:
XX APPLICATION NUMBER: US/07/367,751
XX FILING DATE: 21-JUN-1989
XX PRIOR APPLICATION DATA:
XX APPLICATION NUMBER: 210,594
XX FILING DATE: 23-JUN-1988
XX SEQ ID NO: 4:
XX LENGTH: 168
XX SEQUENCE 182 AA; 19707 MW; 182144 CN;
SQ
Query Match 71.2%; Score 47; DB 4; Length 168;
Best Local Similarity 66.7%; Pred. No. 6.62e+01;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 12 YLATASTMD 20
QY 3 YLISTASSLD 11
RESULT 6
US-08-327-357A-1 STANDARD; PRT: 170 AA.
xxxxxx
Sequence 1, Application US/08327357A
CC Sequence 1, Application US/08327357A
CC Patent No. 5817629
CC GENERAL INFORMATION:
CC APPLICANT: WARREN, Kenneth G.
CC APPLICANT: CATZ, Ingrid
CC TITLE OF INVENTION: PEPTIDE SPECIFICITY OF ANTI-MYELIN BASIC
CC TITLE OF INVENTION: PROTEIN AND THE ADMINISTRATION OF MYELIN BASIC PROTEIN
CC NUMBER OF SEQUENCES: 1
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Venable, Baetjer, Howard & Civiletti
CC STREET: 1201 New York Avenue, N.W., Suite 1000
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA
CC ZIP: 20005
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/327,357A
CC FILING DATE: 21-OCT-1994
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/798,099
CC FILING DATE: 27-NOV-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: CA 2,053,799-0
CC FILING DATE: 22-OCT-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Ihnen, Jeffrey L.
CC REGISTRATION NUMBER: 28,957
CC REFERENCE/DOCKET NUMBER: 27052-115469
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 202-962-4810
CC TELEFAX: 202-962-8300
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 170 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC ORIGINAL SOURCE:
CC ORGANISM: Homo sapiens
CC IMMEDIATE SOURCE:
CC CLONE: human myelin basic protein
SQ SEQUENCE 170 AA; 18459 MW; 143992 CN;
Query Match 71.2%; Score 47; DB 2; Length 170;
Best Local Similarity 66.7%; Pred. No. 6.62e+01;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 14 YLATASTMD 22
||:||||:|

[illegible]

```

XX Sequence 5, Application US/08457245
CC Sequence 5, Application US/08457245
CC Patent No. 5573915
CC GENERAL INFORMATION:
CC APPLICANT: BARRY III, Clifton E.
CC APPLICANT: YUAN, Ying
CC TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA INVOLVED
CC TITLE OF INVENTION: IN THE BIOSYNTHESIS OF CYCLOPROPANATED MYCOLIC ACIDS IN
CC NUMBER OF SEQUENCES: 21
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC STREET: Stewart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/457,245
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Chambers, Guy W.
CC REGISTRATION NUMBER: 30,617
CC REFERENCE/DOCKET NUMBER: 15280-216000
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ. ID NO.: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 384 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: ORF2 protein
CC SEQ. ID NO.: 5
CC SEQUENCE 384 AA; 41963 MW; 701271 CN;
DB 236 SYLPTKALD 245
OY 2 SYLSTASSLD 11
RESULT 9
ID 5194425-3 STANDARD: PRT: 184 AA.
AC xxxxxx
DT 01-JAN-1900
DE Patent No. 5194425.
CC Patent No. 5194425
CC APPLICANT: SHARMA, SOMESH D.; LERCH, L. BERNARD; CLARK,
CC BRIAN R.
CC TITLE OF INVENTION: MHC-MEDIATED TOXIC CONJUGATES USEFUL IN
CC AMELIORATING AUTOIMMUNITY
CC NUMBER OF SEQUENCES: 9
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/367,751
CC FILING DATE: 21-JUN-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 210,594
CC FILING DATE: 23-JUN-1988
CC SEQ. ID NO.: 3:

```


CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/413,118
CC FILING DATE: 29-MAR-1995
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/220,151
CC FILING DATE: 30-MAR-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: FROMMER, WILLIAM S.
CC REGISTRATION NUMBER: 25,506
CC REFERENCE/DOCKET NUMBER: 454310-2670
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 840-3333
CC TELEFAX: (212) 840-0712
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 913 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: N-terminal
CC SEQUENCE 913 AA; 100233 MW; 4073940 CN;
SQ

Query Match 65.2%; Score 43; DB 1; Length 913;
Best Local Similarity 36.4%; Pred. No. 1.81e+02;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Db 862 IRYMSIVSALE 872
QY 1 VSYLSTASSLD 11

RESULT 15
ID US-08-356-786-2 STANDARD; PRT; 1255 AA.
XX
AC xxxxxx
DT
DT
XX
XX
DE
XX
CC Sequence 2, Application US/08356786
CC Patent No. 5877305
CC GENERAL INFORMATION:
CC APPLICANT: Huston, James S.
CC APPLICANT: Oppermann, Hermann
CC APPLICANT: Houston, L. L.
CC APPLICANT: Ring, David B.
CC TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
CC TITLE OF INVENTION: Marker
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
CC STREET: Exchange Place, 53 State Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/356,786
CC FILING DATE:
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: 07/831,967
CC FILING DATE: 06-FEB-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Pitcher, Edmund R.
CC REGISTRATION NUMBER: 27,829
CC REFERENCE/DOCKET NUMBER: CRP-053
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 248-7100
CC TELEFAX: (617) 248-7100
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1255 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1255 AA; 137909 MW; 811405 CN;
SQ

Query Match 65.2%; Score 43; DB 2; Length 1255;
Best Local Similarity 50.0%; Pred. No. 1.81e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 62 LTVLPTNASL 71
QY 1 VSYLSTASSL 10

Search completed: Thu Sep 2 11:22:39 1999
Job time : 7 secs.

THIS PAGE BLANK (USPTO)

 W E S T
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msearch - protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Sep 2 11:18:01 1999; Maspar time 2.14 Seconds

 Molecular output not generated. 118.732 Million cell updates/sec

Title: >US-08-599-226-3
 Description: (1-9) from US08599226.pep
 Perfect Score: 66
 Sequence: 1 QRYNAPYX 9

Scoring table: PAM 150
 Gap 15

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 08
 Listing first 45 summaries

Database: swiss-prot
 I:swissprot

Statistics: Mean 23.600; Variance 27.313; scale 0.864

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	57	86.4	279	1	PRC2_DROME	8.67e-02
2	54	81.8	263	1	PRC2_RAT	4.20e-01
3	54	81.8	263	1	PRC2_HUMAN	4.20e-01
4	51	77.3	1262	1	GNRP_MOUSE	1.92e+00
5	50	75.8	287	1	FCAR_HUMAN	3.15e+00
6	49	74.2	360	1	DCAM_SOLTU	5.13e+00
7	48	72.7	285	1	RS2_ARATH	8.28e+00
8	47	71.2	14	1	HT14_PIG	1.33e+01
9	47	71.2	288	1	TRAB_NEIME	1.33e+01
10	47	71.2	323	1	GDIA_MOUSE	1.33e+01
11	47	71.2	442	1	VIT2_DROME	1.33e+01
12	47	71.2	445	1	GDIB_HUMAN	1.33e+01
13	47	71.2	445	1	GDIB_MOUSE	1.33e+01
14	47	71.2	445	1	GDIB_RAT	1.33e+01
15	47	71.2	447	1	GDIA_HUMAN	1.33e+01
16	47	71.2	447	1	GDIA_BOVIN	1.33e+01
17	47	71.2	451	1	GDIL_YEAST	1.33e+01
18	47	71.2	451	1	GDIL_MOUSE	1.33e+01
19	47	71.2	689	1	YBP3_YEAST	1.33e+01
20	47	71.2	1320	1	PUTA_ECOLI	1.33e+01
21	47	71.2	1320	1	PUTA_SALTY	1.33e+01
22	47	71.2	3305	1	APLP_MANSE	1.33e+01
23	46	69.7	182	1	HPRT_ECOLI	2.11e+01

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
24	46	69.7	360	1	FARX_LYMS	2.11e+01
25	46	69.7	365	1	NADA_SALTY	2.11e+01
26	46	69.7	610	1	EDD_PSEAE	2.11e+01
27	46	69.7	1385	1	YMS5_CABEL	2.11e+01
28	45	68.2	108	1	KVIL_HUMAN	3.33e+01
29	45	68.2	255	1	YAFU_ECOLI	3.33e+01
30	45	68.2	406	1	BHMT_HUMAN	3.33e+01
31	45	68.2	407	1	BHMT_MOUSE	3.33e+01
32	45	68.2	423	1	TRPB_HALVO	3.33e+01
33	45	68.2	467	1	MATK_RAT	3.33e+01
34	45	68.2	491	1	CG2A_DROME	3.33e+01
35	45	68.2	494	1	URIDINE 5'-MONOPHOSPH	3.33e+01
36	45	68.2	505	1	PYR5_NAEGR	3.33e+01
37	45	68.2	507	1	MATK_MOUSE	3.33e+01
38	45	68.2	521	1	MATK_HUMAN	3.33e+01
39	45	68.2	563	1	NPRE_BACAM	3.33e+01
40	45	68.2	563	1	LIP2_GECN	3.33e+01
41	45	68.2	563	1	LIP1_GECN	3.33e+01
42	45	68.2	2410	1	HSCA_ECOLI	3.33e+01
43	45	68.2	2412	1	POL1_BAYMG	3.33e+01
44	45	68.2	245	1	POL1_BAYMG	3.33e+01
45	44	66.7	615	1	Y382_MERJA	5.22e+01
46	44	66.7	615	1	ALBU_CHICK	5.22e+01

ALIGNMENTS

RESULT 1
 ID PRC2_DROME STANDARD; PRT; 279 AA.
 AC P12881;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE PROTEASOME 29 KD SUBUNIT (EC 3.4.99.46) (MULTICATALYTIC ENDOPEPTIDASE DE COMPLEX 35 KD SUBUNIT).
 GN PROS35 OR PROS-35.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
 OC DROSOPHILINAE; DROSOPHILA.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 4-18 AND 194-206.
 RC STRAIN-OREGON-R;
 RX MEDLINE; 90005444.
 RA HAASS C., PESOLD-HURT B., MOLTNAUP G., BEYREUTHER K.,
 RA KLOETZEL P.-M.;
 RT "The PROS-35 gene encodes the 35 kd protein subunit of Drosophila melanogaster proteasome";
 RL EMBO J. 8:2373-2379(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OREGON-R;
 RX MEDLINE; 92249308.
 RA FRENZEL S., TROXEL M., HAASS C., PESOLD-HURT B., GLAETZER K.H.,
 RA KLOETZEL P.-M.;
 RT "Molecular characterization of the genomic regions of the Drosophila alpha-type subunit proteasome genes PROS-dm28.1 and PROS-dm35";
 RL EUR. J. BIOCHEM. 205:1043-1051(1992).
 CC - FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT NEUTRAL OR SLIGHTLY BASIC PH. THE 35 KD SUBUNIT IS PROBABLY A REGULATORY SUBUNIT. THE PROTEASOME HAS AN ATP-DEPENDENT PROTEOLYTIC ACTIVITY.
 CC - PATHWAY: IS INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL PROTEOLYTIC PATHWAY.
 CC - SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE.
 CC - SUBCELLULAR LOCATION: PROTEASOMES ARE FOUND IN THE CYTOPLASM AND ALSO IN THE NUCLEUS.
 CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1A; ALSO KNOWN AS THE PROTEASOME; A-TYPE FAMILY. BELONGS TO THE C2 SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15497; G8382; -
DR EMBL; X62285; G8388; -
DR PIR; S05507; SNFF5K; -
DR PIR; S23450; S23450; -
DR FLYBASE; FBgn0003151; Pro35.
DR PROSITE; PS00388; PROTEASOME_A; 1.
DR PFAM; PF00227; Proteasome; 1.
DR HSSP; P25156; 1PMA.
KW PROTEASOME; HYDROLASE; PROTEASE; PHOSPHORYLATION.
FT MOD_RES 103 103 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 279 AA; 31058 MW; AA18BBD6 CRC32;

Query Match 86.4%; Score 57; DB 1; Length 279;
Best Local Similarity 75.0%; Pred. No. 8,67e-02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 121 ORYDPRY 128
|||:|:|
QY 1 ORYDPRY 8

RESULT 2 STANDARD; PRT; 263 AA.
ID PRC2_RAT
AC P18420;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROTEASOME COMPONENT C2 (EC 3.4.99.46) (MACROPAIN SUBUNIT C2)
DE (PROTEASOME NU CHAIN) (MULTICATALYTIC ENDOPEPTIDASE COMPLEX SUBUNIT
DE C2).
GN PSMAL.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ROSENTIA; SCIUROMATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-LIVER.
RX MEDLINE; 90057428.
RA FUJIWARA T., TANAKA K., KUMATORI A., SHIN S., YOSHIMURA T.,
RA ICHIHARA A., TORUNAGA F., ARUGA R., IWANAGA S., KAKIZUKA A.,
NAKANISHI S.;
"Molecular cloning of cDNA for proteasomes (multicatalytic proteinase
complexes) from rat liver: primary structure of the largest component
(C2).";
RL BIOCHEMISTRY 28:7332-7340(1989).
RN [2]
RP SEQUENCE OF 1-30.
RC TISSUE-LIVER.
RX MEDLINE; 90243011.
RA TORUNAGA F., ARUGA R., IWANAGA S., TANAKA K., ICHIHARA A., TAKAO T.,
RA SHIMONISHI Y.;
"The NH2-terminal residues of rat liver proteasome (multicatalytic
proteinase complex) subunits, C2, C3 and C8, are N
alpha-acetylated";
RL FEBS LETT 263:373-375(1990).
CC -1- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
CC WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH
CC ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT
CC NEUTRAL OR SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT
CC PROTEOLYTIC ACTIVITY.
CC -1- PATHWAY: IS INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL
CC PROTEOLYTIC PATHWAY.
CC -1- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
CC SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE.
CC -1- SUBCELLULAR LOCATION: PROTEASOMES ARE FOUND IN THE CYTOPLASM AND
CC ALSO IN THE NUCLEUS.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL RAT TISSUES.
CC -1- PTM: ITS C-TERMINAL EXTENSION IS PARTIALLY CLEAVED OFF BY LIMITED
CC PROTEOLYSIS LEADING TO A CONVERSION OF THE PROTEASOME FROM ITS
CC LATENT INTO ITS ACTIVE FORM.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1A; ALSO KNOWN AS THE
CC PROTEASOME A-TYPE FAMILY. BELONGS TO THE C2 SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M29859; G206382; -
DR EMBL; D90265; G220877; -
DR PIR; A32968; SNRRC2.
DR PROSITE; PS00388; PROTEASOME_A; 1.
DR PFAM; PF00227; Proteasome; 1.
DR HSSP; P25156; 1PMA.
KW PROTEASOME; HYDROLASE; PROTEASE; ACETYLATION.
FT MOD_RES 1 1 ACETYLATION.
SQ SEQUENCE 263 AA; 29517 MW; 01D55620 CRC32;

Query Match 81.8%; Score 54; DB 1; Length 263;
Best Local Similarity 75.0%; Pred. No. 4,20e-01;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 121 ORYDPRY 128
|||:|:|
QY 1 ORYDPRY 8

RESULT 3 STANDARD; PRT; 263 AA.
ID PRC2_HUMAN
AC P25786;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE PROTEASOME COMPONENT C2 (EC 3.4.99.46) (MACROPAIN SUBUNIT C2)
DE (PROTEASOME NU CHAIN) (MULTICATALYTIC ENDOPEPTIDASE COMPLEX SUBUNIT
DE C2) (30 KD PROSOMAL PROTEIN) (PROS-30).
GN PSMAL OR PSC2 OR PROS30.
GN HOMO SAPIENS (HUMAN).
OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ROSENTIA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93013039.
RA SILVA-PEREIRA I., BEY F., COUX O., SCHERRER K.;
"Two mRNAs exist for the hs PROS-30 gene encoding a component of
human prosomes";
RL GENE 120:235-242(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91223105.
RA TAMURA T., LEE D.H., OSAKA F., FUJIWARA T., SHIN S., CHUNG C.H.,
RA TANAKA K., ICHIHARA A.;
"Molecular cloning and sequence analysis of cDNAs for five major
subunits of human proteasomes (multi-catalytic proteinase
complexes).";
RL BIOCHIM. BIOPHYS. ACTA 1089:95-102(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91363412.
RA DEBARTINO G.N., ORTH K., MCCULLOUGH M.L., LEE L.W., MUNN T.Z.,
RA MOOMAN C.R., DAWSON P.A., SLAUGHTER C.A.;
"The primary structures of four subunits of the human,
high-molecular-weight proteinase, macropain (proteasome), are
distinct but homologous";
RL BIOCHIM. BIOPHYS. ACTA 1079:29-38(1991).
CC -1- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX

CC WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH
 CC ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT
 CC NEUTRAL OR SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT
 CC PROTEOLYTIC ACTIVITY.
 CC -1- PATHWAY: IS INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL
 CC PROTEOLYTIC PATHWAY.
 CC -1- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
 CC SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE.
 CC -1- SUBCELLULAR LOCATION: PROTEASOMES ARE FOUND IN THE CYTOPLASM AND
 CC ALSO IN THE NUCLEUS.
 CC -1- ALTERNATIVE PRODUCTS: TWO FORMS (SHORT AND LONG) ARE PRODUCED BY
 CC ALTERNATIVE SPLICING OF THE GENE FOR THIS PROTEIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1A; ALSO KNOWN AS THE
 CC PROTEASOME A-TYPE FAMILY. BELONGS TO THE C2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M64992; GI90047; ALT_SEQ.
 CC EMBL: D00759; G220022; -.
 CC EMBL: X61969; G296738; -.
 CC PIR: S15897; S15897.
 CC PIR: J01445; J01445.
 CC AARHUS/GENET-2DPAGE; 2223; IEF.
 CC MIM: 602854; -.
 CC DR PROSITE: PS00388; PROTEASOME_A; 1.
 CC DR PFAM: PF00227; proteasome; 1.
 CC DR HSSP: P25156; 1PMA.
 CC KW PROTEASOME; HYDROLASE; PROTEASE; ACETYLATION; ALTERNATIVE SPLICING.
 CC FT MOD.RES 1 1 ACETYLATION (BY SIMILARITY).
 CC FT VARSPLIC 1 1 M->MOLSKV (IN LONG FORM).
 CC SQ SEQUENCE 263 AA; 29555 MW; 6CD09A93 CRC32;
 CC -----
 CC Query Match 81.8%; Score 54; DB 1; Length 263;
 CC Best Local Similarity 75.0%; Pred. No. 4.20e-01;
 CC Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 CC Db 121 ORYGRAPY 128
 CC 11111111
 CC QY 1 ORYGRAPY 8
 CC -----
 CC RESULT 4 STANDARD: PRT; 1262 AA.
 CC GNR_MOUSE
 CC P27671;
 CC 01-AUG-1992 (REL. 23, CREATED)
 CC 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 CC 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 CC GUANINE NUCLEOTIDE RELEASING PROTEIN (GNRP).
 CC RASGRF1 OR CDC25MM OR GRF1.
 CC OS MUS MUSCULUS (MOUSE).
 CC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
 CC RODENTIA: SCIROGNATHI: MURIDAE: MURINAE: MUS.
 CC RN [1]
 CC RC SEQUENCE FROM N.A.
 CC RC STRAIN-BALB/C;
 CC RX MEDLINE: 93010996.
 CC CEN H. LOWY D.D.;
 CC RA "Isolation of multiple mouse cDNAs with coding homology to
 CC RT Saccharomyces cerevisiae CDC25: Identification of a region related to
 CC RT Bcr, Vav, Dbl and CDC24.";
 CC RL EMBO J. 11:4007-4015(1992).
 CC RN [2]
 CC RP SEQUENCE OF 791-1262 FROM N.A.
 CC RC STRAIN-SWISS: TISSUE-BRAIN;
 CC RX MEDLINE: 92289680.
 CC RA MARTEGANI E., VANONI M., ZIPPEL R., COCCETTI P., BRAMBILLA R.,
 CC FERRARI C., STURANI E.P., ALBERGHINA L.;

RT "Cloning by functional complementation of a mouse cDNA encoding a
 RT homologue of CDC25, a Saccharomyces cerevisiae RAS activator.";
 RT EMBO J. 11:2151-2157(1992).
 RN [3]
 RP SEQUENCE OF 1031-1226 FROM N.A.
 RX MEDLINE: 9235779.
 RA WEI W., MOSTELLER R.D., SANYAL P., GONZALES E., MCKINNEY D.,
 RA DASGUPTA C., LI P., LIU B.X., BROEK D.;
 RT "Identification of a mammalian gene structurally and functionally
 RT related to the CDC25 gene of Saccharomyces cerevisiae.";
 RL PROC. NATL. ACAD. SCI. U.S.A. 89:7100-7104(1992).
 CC -1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: TO OTHER GUANINE-NUCLEOTIDE RELEASING FACTORS OF THE
 CC CDC25 FAMILY.
 CC -----
 CC -1- SIMILARITY: CONTAINS 2 PH DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L20899; G388247; -.
 CC EMBL: X59868; G50358; -.
 CC PIR: S20730; S20730.
 CC PIR: S20730; S20730.
 CC PIR: S20730; S20730.
 CC MGD: MGI:99694; RASGRF1.
 CC DR PROSITE: PS00720; GDS_CDC25; 1.
 CC DR PROSITE: PS00741; GDS_CDC24; 1.
 CC DR PROSITE: PS50003; PH_DOMAIN; 2.
 CC DR PFAM: PF00169; PH; 2.
 CC DR PFAM: PF00612; IQ; 1.
 CC DR PFAM: PF00617; RASGEF; 1.
 CC DR PFAM: PF00618; RASGEF; 1.
 CC DR PFAM: PF00621; RHOGEF; 1.
 CC KW GUANINE-NUCLEOTIDE RELEASING FACTOR.
 CC FT DOMAIN 22 130 PH.
 CC FT DOMAIN 460 588 PH.
 CC FT CONFLICT 1033 1033 E->D (IN REF. 3).
 CC SQ SEQUENCE 1262 AA; 144101 MW; 021C787F CRC32;
 CC -----
 CC Query Match 77.3%; Score 51; DB 1; Length 1262;
 CC Best Local Similarity 50.0%; Pred. No. 1.92e+00;
 CC Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 CC Db 1060 EKERPTY 1067
 CC 11111111
 CC QY 1 ORYGRAPY 8
 CC -----
 CC RESULT 5 STANDARD: PRT; 287 AA.
 CC FCAR_HUMAN
 CC P24071: 015728; 013603; 013604;
 CC 01-MAR-1992 (REL. 21, CREATED)
 CC 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 CC 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 CC DE IMMUNOGLOBULIN ALPHA FC RECEPTOR PRECURSOR (IGA FC RECEPTOR) (CD89
 CC ANTIGEN).
 CC GN FCAR OR CD89.
 CC OS HOMO SAPIENS (HUMAN).
 CC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
 CC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE: 91079769.
 CC RA MALISZEWSKI C.R., MARCH C.J., SCHOENBORN M.A., GIMPEL S., SHEN L.;
 CC "Expression cloning of a human Fc receptor for Iga.";
 CC J. EXP. MED. 172:1665-1672(1990).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.

CC AND ROOTS.
 CC -I- DEVELOPMENTAL STAGE: TRANSCRIBED IN THE STOLON TIP DURING THE
 CC EARLY STAGES OF TUBERIZATION. MAXIMUM EXPRESSION WAS IN NON-
 CC SMELLING STOLON TIPS FROM STAGE B, AND LEVEL DECLINED AS THE
 CC TUBER INCREASED IN SIZE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: 211680; G21485; -
 CC DR EMBL: 574514; G807094; -
 CC DR PIR: S28047; S28047; -
 CC KW SPERMIDINE BIOSYNTHESIS; LYSASE; DECARBOXYLASE; PYRUVATE; ZMOGEN.
 CC CHAIN 1 72
 CC CHAIN (BY SIMILARITY).
 CC S-ADENOSYLMETHIONINE DECARBOXYLASE BETA
 CC CHAIN (BY SIMILARITY).
 CC CLEAVAGE (NONHYDROLITICAL)
 CC (BY SIMILARITY).
 CC CONVERTED TO A PYRUVYL GROUP
 CC (BY SIMILARITY).
 CC IMPORTANT FOR CATALYTIC ACTIVITY (BY
 CC SIMILARITY).
 CC IMPORTANT FOR CATALYTIC ACTIVITY (BY
 CC SIMILARITY).
 CC ACT_SITE 16 16
 CC FT ACT_SITE 16 16
 CC FT ACT_SITE 87 87
 CC FT ACT_SITE 87 87
 CC FT ACT_SITE 174 174
 CC FT CONFLICT 257 257
 CC FT CONFLICT 291 291
 CC FT CONFLICT 305 305
 CC FT CONFLICT 305 305
 CC SQ SEQUENCE 360 AA: 39726 MW: 4F29EFEE CRC32:
 CC
 CC Query Match 74.2%; Score 49; DB 1; Length 360;
 CC Best Local Similarity 50.0%; Pred. No. 5.13e+00;
 CC Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC Db 332 OKFRTPTV 339
 CC Y 1 OKRYNRPV 8
 CC
 CC RESULT 7
 CC RS2_ARATH STANDARD; PRT; 285 AA.
 CC P49688: 022936;
 CC 01-FEB-1996 (REL. 33, CREATED)
 CC 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
 CC 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 CC 40S RIBOSOMAL PROTEIN S2.
 CC RPS2 OR T11A07.6.
 CC GN
 CC OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
 CC EUKARYOTA: VIRIDIPHYTES: STREPTOPHYTA: EMBRYOPHYTA: TRACHEOPHYTA;
 CC EUPHYLOPHYTES: SPERMATOPHYTA: MAGNOLIOPHYTA: EUDICOTYLEDONS; ROSIDAE;
 CC OC CAPRARALES: BRASSICACEAE; ARABIDOPSIS.
 CC OC
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-CV. COLUMBIA;
 CC ROUNLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,
 CC SPRIGGS T.A., MASON T.M., KERLAVAGE A.R., ADAMS M.D., SOMERVILLE C.R.,
 CC VENTER J.C.;
 CC RA SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC RN [2]
 CC RP SEQUENCE OF 119-285 FROM N.A.
 CC RC STRAIN-CV. COLUMBIA;
 CC RAYNAL M., GRELET F., LAUDIE M., MEYER Y., COOKE R., DELSENY M.;
 CC SUBMITTED (OCT-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC CC -I- SIMILARITY: BELONGS TO THE SSP FAMILY OF RIBOSOMAL PROTEINS.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AC002339; G2335095; -
 CC DR EMBL: 217622; G16781; -
 CC DR PROSITE: PS00585; RIBOSOMAL_S5; 1.
 CC DR PFAM: PF00333; S5; 1.
 CC DR HSSP: P02357; LPKP.
 CC KW RIBOSOMAL PROTEIN.
 CC FT CONFLICT 268 269 AS -> ST (IN REF. 2).
 CC FT CONFLICT 272 281 LSTSKRPDPTV -> VSATVITTEG (IN REF. 2).
 CC SQ SEQUENCE 285 AA: 30878 MW: 4684B4BF CRC32:
 CC
 CC Query Match 72.7%; Score 48; DB 1; Length 285;
 CC Best Local Similarity 57.1%; Pred. No. 8.28e+00;
 CC Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Db 254 RERSRP 260
 CC Y 2 RYRNRPV 8
 CC
 CC RESULT 8
 CC ID HY14_PIG STANDARD; PRT; 14 AA.
 CC AC P01155;
 CC DT 21-JUL-1986 (REL. 01, CREATED)
 CC DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 CC DT 21-JUL-1986 (REL. 01, LAST ANNOTATION UPDATE)
 CC DE HYPOTHALAMIC TETRADECAPEPTIDE.
 CC SOS SCROFA (PIG).
 CC OS EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA;
 CC OC ARTIODACTYLA: SUIFORMES; SUINA; SUIDAE; SUS.
 CC RN [1]
 CC RP SEQUENCE.
 CC RA SCHLESINGER D.H., NIALL H.D., LINTHICUM G.L., DUPONT A.,
 CC RA SCHALLY A.V.;
 CC RL SUBMITTED (NOV-1976) TO THE PIR DATA BANK.
 CC DR PIR: A01419; NYPG14.
 CC DR HSSP: P21856; IGND.
 CC KW AMIDATION.
 CC FT MOD_RES 14 14
 CC FT MOD_RES 14 14
 CC SQ SEQUENCE 14 AA: 59F0BC0B CRC32:
 CC
 CC Query Match 71.2%; Score 47; DB 1; Length 14;
 CC Best Local Similarity 57.1%; Pred. No. 1.33e+01;
 CC Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC Db 2 RYKSPV 8
 CC Y 2 RYRNRPV 8
 CC
 CC RESULT 9
 CC ID TRAG_NEIME STANDARD; PRT; 288 AA.
 CC AC 000840;
 CC DT 01-JUN-1994 (REL. 29, CREATED)
 CC DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 CC DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
 CC DE TRANSPOSASE FOR INSERTION SEQUENCE ELEMENT IS1106 (ORF 1).
 CC OS NEISSERIA MENINGITIDIS.
 CC OC BACTERIA: PROTEOBACTERIA: BETA SUBDIVISION; NEISSERIACEAE; NEISSERIA.
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-B15;
 CC RX MEDLINE: 92326641.
 CC RA KNIGHT A.I., NI H., CARTWRIGHT K.A.V., MCFADDEN J.;
 CC RT "Identification and characterization of a novel insertion sequence,
 CC IS1106, downstream of the porA gene in B15 Neisseria meningitidis";

```

RL  MOL. MICROBIOL. 6:1565-1573(1992).
CC  -1- FUNCTION: INVOLVED IN THE TRANSDUCTION OF THE INSERTION
CC  SEQUENCE.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  DR  EMBL: Z11857; G45066; -.
CC  DR  PIR: S22628; S22628.
CC  KW  TRANSPOSABLE ELEMENT; TRANSDUCTION; DNA-BINDING; DNA RECOMBINATION.
CC  SQ  SEQUENCE 288 AA; 32758 MW; 068FE8FA CRC32;

Query Match 71.2%; Score 47; DB 1; Length 288;
Best Local Similarity 71.4%; Pred. No. 1.33e+01;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 250 RYRABAY 256
QY 2 RYRABAY 8

RESULT 10
ID  GDI1_MOUSE STANDARD: PRT; 323 AA.
AC  P50396;
DT  01-OCT-1996 (REL. 34, CREATED)
DT  01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT  15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE  RAB GDP DISSOCIATION INHIBITOR ALPHA (RAB GDI ALPHA) (GDI-1)
DE  (FRAGMENT).
DE  GDI1 OR RABGDI1.
OS  MUS MUSCULUS (MOUSE).
OC  EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC  RODENTIA; SCURIONATHI; MORIDAE; MORINAE; MUS.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-BALB/C; TISSUE-SKELETAL MUSCLE;
RX  MEDLINE: 94217740.
RA  SHISHOVA A., SUDHOFF T.C., CZECH M.P.;
RT  "Cloning, characterization, and expression of a novel GDP
RT  dissociation inhibitor isoform from skeletal muscle.";
RL  MOL. CELL. BIOL. 14:3459-3468(1994).
CC  -1- FUNCTION: REGULATES THE GDP/GTP EXCHANGE REACTION OF MOST RAB
CC  PROTEINS BY INHIBITING THE DISSOCIATION OF GDP FROM THEM, AND THE
CC  SUBSEQUENT BINDING OF GTP TO THEM.
CC  -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC  -1- TISSUE SPECIFICITY: HIGH EXPRESSION IN BRAIN, LOWER IN OTHER
CC  TISSUES.
CC  -1- SIMILARITY: BELONGS TO THE TCD/MRS6 FAMILY OF GDP DISSOCIATION
CC  INHIBITOR.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  DR  EMBL: U07950; G516537; -.
CC  DR  MGD: MGI:99846; GDI1.
CC  DR  PFM: PFO0996; GDI: 1.
CC  DR  HSP: P21856; 1GND.
CC  KW  GTPASE ACTIVATION.
CC  FT  NON_TER 1
CC  SQ  SEQUENCE 323 AA; 36601 MW; EF8EC281 CRC32;

Query Match 71.2%; Score 47; DB 1; Length 323;
Best Local Similarity 57.1%; Pred. No. 1.33e+01;

```

```

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 94 RYKSPY 100
QY 2 RYRABAY 8

RESULT 11
ID  VIT2_DROME STANDARD: PRT; 442 AA.
AC  P02844;
DT  21-JUL-1986 (REL. 01, CREATED)
DT  21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT  01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE  VITELLOGENIN II PRECURSOR (YOLK PROTEIN 2).
DE  YP2.
OS  DROSOPHILA MELANOGASTER (FRUIT FLY).
OC  EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC  PTERYOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYROIDERA;
OC  DROSOPHILIDAE; DROSOPHILA.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE: 83189120.
RA  HUNG M.-C., WENSINK P.C.;
RT  "Sequence and structure conservation in yolk proteins and their
RT  genes.";
RL  J. MOL. BIOL. 164:481-492(1983).
CC  -1- FUNCTION: VITELLOGENIN IS THE MAJOR YOLK PROTEIN OF EGGS WHERE
CC  IT IS USED AS A FOOD SOURCE DURING EMBRYOGENESIS.
CC  -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE FAT BODY AND OVARIAN
CC  FOLLICLE CELLS AND ACCUMULATE IN THE OOCYTE.
CC  -1- INDUCTION: IN MALES BY BETA-ECDSONE.
CC  -1- SIMILARITY: PARTIAL, TO LIPASES. STRONG TO OTHER VITELLOGENINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  DR  EMBL: J01157; -. NOT_ANNOTATED_CDS.
CC  DR  PIR: A03333; VITF2.
CC  DR  FLYBASE: FBgn0005391; YP2.
CC  DR  PFM: PFO0151; 1lpase; 1.
CC  KW  YOLK; SIGNAL.
CC  FT  SIGNAL 1 20
CC  FT  CHAIN 21 442 VITELLOGENIN II.
CC  SQ  SEQUENCE 442 AA; 45678 MW; A43872E CRC32;

Query Match 71.2%; Score 47; DB 1; Length 442;
Best Local Similarity 75.0%; Pred. No. 1.33e+01;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 160 QRYNOPY 167
QY 1 QRYNABAY 8

RESULT 12
ID  GDI1_HUMAN STANDARD: PRT; 445 AA.
AC  P50395;
DT  01-OCT-1996 (REL. 34, CREATED)
DT  01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT  15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE  RAB GDP DISSOCIATION INHIBITOR BETA (RAB GDI BETA) (GDI-2).
DE  GDI2 OR RABGDI2.
OS  HOMO SAPIENS (HUMAN).
OC  EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC  PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE-BRAIN;

```

RA ASADA M., KAIBUCHI K., TAKAI Y.:
RL SUBMITTED (APR-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
CC [2]
CC TISSUE SPECIFICITY.
RX MEDLINE: 95359978.
RA BACHNER D., SELACKER Z., KORN B., HAMEISTER H., POUSTKA A.:
RT "Expression patterns of two human genes coding for different rab GDP-
RT dissociation inhibitors (Gdis), extremely conserved proteins involved
RT in cellular transport."
RL HUM. MOL. GENET. 4:701-708(1995).
CC -1- FUNCTION: REGULATES THE GDP/GTP EXCHANGE REACTION OF MOST RAB
CC PROTEINS BY INHIBITING THE DISSOCIATION OF GDP FROM THEM, AND THE
CC SUBSEQUENT BINDING OF GTP TO THEM.
CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
CC -1- SIMILARITY: BELONGS TO THE TCD/MRS6 FAMILY OF GDP DISSOCIATION
CC INHIBITOR.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: D13988; G285975; -
CC MIM: 600767; -
CC DR PFAM: PF00996; GDI: 1.
CC DR HSSP: P21856; 1GND.
CC KW GTPASE ACTIVATION.
CC SQ SEQUENCE 445 AA; 50664 MW; 858E540A CRC32;
Query Match 71.2%; Score 47; DB 1; Length 445;
Best Local Similarity 57.1%; Pred. No. 1.33e+01;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 218 RYKSPY 224
QY 2 RYNNAPY 8
RESULT 13
ID GDIB_RAT STANDARD; PRT; 445 AA.
AC P50399;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE RAB GDP DISSOCIATION INHIBITOR BETA (RAB GDI BETA) (GDI-2).
CC GDI2 OR RABGDI2.
CC RATIUS NORVEGICUS (RAT).
CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
CC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;
CC MEDLINE: 94245743.
CC RA NISHIMURA N., NAKAMURA H., TAKAI Y., SANO K.:
RT "Molecular cloning and characterization of two rab GDI species from
RT rat brain: brain-specific and ubiquitous types."
RL J. BIOL. CHEM. 269:14191-14198(1994).
CC [2]
CC SEQUENCE OF 30-54 AND 58-74, AND CHARACTERIZATION.
CC MEDLINE: 95298038.
CC RA ARAKI K., NAKAMISHI H., HIRANO H., KATO M., SASAKI T., TAKAI Y.:
RT "Purification and characterization of Rab GDI beta from rat brain."
CC BIOCHEM. BIOPHYS. RES. COMMUN. 211:296-305(1995).
CC -1- FUNCTION: REGULATES THE GDP/GTP EXCHANGE REACTION OF MOST RAB
CC PROTEINS BY INHIBITING THE DISSOCIATION OF GDP FROM THEM, AND THE
CC SUBSEQUENT BINDING OF GTP TO THEM.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED.
CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED.
CC -1- SIMILARITY: BELONGS TO THE TCD/MRS6 FAMILY OF GDP DISSOCIATION
CC INHIBITOR.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X74401; G396433; -
CC DR PFAM: PF00996; GDI: 1.
CC DR HSSP: P21856; 1GND.
CC KW GTPASE ACTIVATION.
CC SQ SEQUENCE 445 AA; 50685 MW; 842P64F7 CRC32;
Query Match 71.2%; Score 47; DB 1; Length 445;
Best Local Similarity 57.1%; Pred. No. 1.33e+01;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 218 RYKSPY 224
QY 2 RYNNAPY 8
RESULT 14
ID GDIB_MOUSE STANDARD; PRT; 445 AA.
AC P50397;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE RAB GDP DISSOCIATION INHIBITOR BETA (RAB GDI BETA) (GDI-2).
CC GDI2 OR RABGDI2.
CC OS MUS MUSCULUS (MOUSE).
CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
CC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN-BALB/C; TISSUE-SKELETAL MUSCLE;
CC MEDLINE: 94217740.
CC RX SHISHNEVA A., SUDHOR T.C., CZECH M.P.:
RT "Cloning, characterization, and expression of a novel GDP
RT dissociation inhibitor isoform from skeletal muscle."
RL MOL. CELL. BIOL. 14:3459-3468(1994).
CC -1- FUNCTION: REGULATES THE GDP/GTP EXCHANGE REACTION OF MOST RAB
CC PROTEINS BY INHIBITING THE DISSOCIATION OF GDP FROM THEM, AND THE
CC SUBSEQUENT BINDING OF GTP TO THEM.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED.
CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED.
CC -1- SIMILARITY: BELONGS TO THE TCD/MRS6 FAMILY OF GDP DISSOCIATION
CC INHIBITOR.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U07951; G1486480; ALT_SEQ.
CC DR MGD; MGI:99845; GDI2.
CC DR PFAM: PF00996; GDI: 1.
CC DR HSSP: P21856; 1GND.
CC KW GTPASE ACTIVATION.
CC SQ SEQUENCE 445 AA; 50653 MW; 79A76704 CRC32;
Query Match 71.2%; Score 47; DB 1; Length 445;
Best Local Similarity 57.1%; Pred. No. 1.33e+01;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 218 RYKSPY 224
QY 2 RYNNAPY 8

RESULT 15
ID GDI1.HUMAN STANDARD: PRT: 447 AA.
AC P31150; P50394;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE RAB GDP DISSOCIATION INHIBITOR ALPHA (RAB GDI ALPHA) (GDI-1) (XAP-4).
GN GDI1 OR RABGDI1 OR XAP4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RN SEQUENCE FROM N.A.
RA SEDLACEK Z., KONECKI D.S., KORN B., KLAUCK S.M., POUSTKA A.;
RT "Evolutionary conservation and genomic organization of XAP-4, an Xq28
located gene coding for a human rab GDP-dissociation inhibitor
(GDI).";
MAMM. GENOME 5:633-639(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE-RETINA;
RX MEDLINE; 96062207.
RA NISHIMURA N., GOTT J., NAKAMURA H., ORITA S., TAKAI Y., SANO K.;
RT "Cloning of a brain-type isoform of human Rab GDI and its expression
in human neuroblastoma cell lines and tumor specimens.";
CANCER RES. 55:5445-5450(1995).
RN [3]
RN SEQUENCE FROM N.A.
RA CHEN E.Y., ZOLLO M., MAZZARELLA R.A., CICCOCICOLA A., CHEN C.N.,
RA ZOU L., HEINER C., BURROGH F.W., RIBETTO M., SCHLESSINGER D.,
RA D'URSO M.;
RN [4]
RN SUBMITTED (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]
RN SEQUENCE OF 143-181 FROM N.A.
RA HOCHGESCHWENDER U.;
RN SUBMITTED (SEP-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [6]
RN SEQUENCE OF 328-436 FROM N.A.
RA BHAT K.S.;
RN SUBMITTED (XXX-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [7]
RN SEQUENCE OF 349-361.
RC TISSUE-KERATINOCYTES;
RX MEDLINE; 93162043.
RA RASMUSSEN H.H., VAN DAMME J., PUYE M., GESSER B., CELIS J.E.,
VANDERKERCKHOVE J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
protein database of normal human epidermal keratinocytes.";
ELECTROPHORESIS 13:960-969(1992).
RN [8]
RN TISSUE SPECIFICITY.
RX MEDLINE; 95359978.
RA BACHNER D., SEDLACEK Z., KORN B., HAMEISTER H., POUSTKA A.;
RT "Expression patterns of two human genes coding for different rab GDP-
dissociation inhibitors (GDIs), extremely conserved proteins involved
in cellular transport.";
HUM. MOL. GENET. 4:701-708(1995).
RN [9]
RN FUNCTION: REGULATES THE GDP/GTP EXCHANGE REACTION OF MOST RAB
PROTEINS BY INHIBITING THE DISSOCIATION OF GDP FROM THEM, AND THE
SUBSEQUENT BINDING OF GTP TO THEM.
RN [10]
RN SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
RN [11]
RN TISSUE SPECIFICITY: BRAIN, PREDOMINANT IN NEURAL AND SENSORY
TISSUES
RN [12]
RN SIMILARITY: BELONGS TO THE TCD/MRS6 FAMILY OF GDP DISSOCIATION
INHIBITOR.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X79354; 6695523; -;
DR EMBL; X79355; 6695523; JOINED.
DR EMBL; X79356; 6695523; JOINED.
DR EMBL; X79357; 6695523; JOINED.
DR EMBL; X79358; 6695523; JOINED.
DR EMBL; X79359; 6695523; JOINED.
DR EMBL; X79360; 6695523; JOINED.
DR EMBL; X79361; 6695523; JOINED.
DR EMBL; X79362; 6695523; JOINED.
DR EMBL; X79363; 6695523; JOINED.
DR EMBL; X79364; 6695523; JOINED.
DR EMBL; X79365; 6695523; JOINED.
DR EMBL; X79366; 6695523; JOINED.
DR EMBL; X79367; 6695523; JOINED.
DR EMBL; X79368; 6695523; JOINED.
DR EMBL; X79369; 6695523; JOINED.
DR EMBL; X79370; 6695523; JOINED.
DR EMBL; X79371; 6695523; JOINED.
DR EMBL; X79372; 6695523; JOINED.
DR EMBL; X79373; 6695523; JOINED.
DR EMBL; X79374; 6695523; JOINED.
DR EMBL; X79375; 6695523; JOINED.
DR EMBL; X79376; 6695523; JOINED.
DR EMBL; X79377; 6695523; JOINED.
DR EMBL; X79378; 6695523; JOINED.
DR EMBL; X79379; 6695523; JOINED.
DR EMBL; X79380; 6695523; JOINED.
DR EMBL; X79381; 6695523; JOINED.
DR EMBL; X79382; 6695523; JOINED.
DR EMBL; X79383; 6695523; JOINED.
DR EMBL; X79384; 6695523; JOINED.
DR EMBL; X79385; 6695523; JOINED.
DR EMBL; X79386; 6695523; JOINED.
DR EMBL; X79387; 6695523; JOINED.
DR EMBL; X79388; 6695523; JOINED.
DR EMBL; X79389; 6695523; JOINED.
DR EMBL; X79390; 6695523; JOINED.
DR EMBL; X79391; 6695523; JOINED.
DR EMBL; X79392; 6695523; JOINED.
DR EMBL; X79393; 6695523; JOINED.
DR EMBL; X79394; 6695523; JOINED.
DR EMBL; X79395; 6695523; JOINED.
DR EMBL; X79396; 6695523; JOINED.
DR EMBL; X79397; 6695523; JOINED.
DR EMBL; X79398; 6695523; JOINED.
DR EMBL; X79399; 6695523; JOINED.
DR EMBL; X79400; 6695523; JOINED.
DR EMBL; X79401; 6695523; JOINED.
DR EMBL; X79402; 6695523; JOINED.
DR EMBL; X79403; 6695523; JOINED.
DR EMBL; X79404; 6695523; JOINED.
DR EMBL; X79405; 6695523; JOINED.
DR EMBL; X79406; 6695523; JOINED.
DR EMBL; X79407; 6695523; JOINED.
DR EMBL; X79408; 6695523; JOINED.
DR EMBL; X79409; 6695523; JOINED.
DR EMBL; X79410; 6695523; JOINED.
DR EMBL; X79411; 6695523; JOINED.
DR EMBL; X79412; 6695523; JOINED.
DR EMBL; X79413; 6695523; JOINED.
DR EMBL; X79414; 6695523; JOINED.
DR EMBL; X79415; 6695523; JOINED.
DR EMBL; X79416; 6695523; JOINED.
DR EMBL; X79417; 6695523; JOINED.
DR EMBL; X79418; 6695523; JOINED.
DR EMBL; X79419; 6695523; JOINED.
DR EMBL; X79420; 6695523; JOINED.
DR EMBL; X79421; 6695523; JOINED.
DR EMBL; X79422; 6695523; JOINED.
DR EMBL; X79423; 6695523; JOINED.
DR EMBL; X79424; 6695523; JOINED.
DR EMBL; X79425; 6695523; JOINED.
DR EMBL; X79426; 6695523; JOINED.
DR EMBL; X79427; 6695523; JOINED.
DR EMBL; X79428; 6695523; JOINED.
DR EMBL; X79429; 6695523; JOINED.
DR EMBL; X79430; 6695523; JOINED.
DR EMBL; X79431; 6695523; JOINED.
DR EMBL; X79432; 6695523; JOINED.
DR EMBL; X79433; 6695523; JOINED.
DR EMBL; X79434; 6695523; JOINED.
DR EMBL; X79435; 6695523; JOINED.
DR EMBL; X79436; 6695523; JOINED.
DR EMBL; X79437; 6695523; JOINED.
DR EMBL; X79438; 6695523; JOINED.
DR EMBL; X79439; 6695523; JOINED.
DR EMBL; X79440; 6695523; JOINED.
DR EMBL; X79441; 6695523; JOINED.
DR EMBL; X79442; 6695523; JOINED.
DR EMBL; X79443; 6695523; JOINED.
DR EMBL; X79444; 6695523; JOINED.
DR EMBL; X79445; 6695523; JOINED.
DR EMBL; X79446; 6695523; JOINED.
DR EMBL; X79447; 6695523; JOINED.
DR EMBL; X79448; 6695523; JOINED.
DR EMBL; X79449; 6695523; JOINED.
DR EMBL; X79450; 6695523; JOINED.
DR EMBL; X79451; 6695523; JOINED.
DR EMBL; X79452; 6695523; JOINED.
DR EMBL; X79453; 6695523; JOINED.
DR EMBL; X79454; 6695523; JOINED.
DR EMBL; X79455; 6695523; JOINED.
DR EMBL; X79456; 6695523; JOINED.
DR EMBL; X79457; 6695523; JOINED.
DR EMBL; X79458; 6695523; JOINED.
DR EMBL; X79459; 6695523; JOINED.
DR EMBL; X79460; 6695523; JOINED.
DR EMBL; X79461; 6695523; JOINED.
DR EMBL; X79462; 6695523; JOINED.
DR EMBL; X79463; 6695523; JOINED.
DR EMBL; X79464; 6695523; JOINED.
DR EMBL; X79465; 6695523; JOINED.
DR EMBL; X79466; 6695523; JOINED.
DR EMBL; X79467; 6695523; JOINED.
DR EMBL; X79468; 6695523; JOINED.
DR EMBL; X79469; 6695523; JOINED.
DR EMBL; X79470; 6695523; JOINED.
DR EMBL; X79471; 6695523; JOINED.
DR EMBL; X79472; 6695523; JOINED.
DR EMBL; X79473; 6695523; JOINED.
DR EMBL; X79474; 6695523; JOINED.
DR EMBL; X79475; 6695523; JOINED.
DR EMBL; X79476; 6695523; JOINED.
DR EMBL; X79477; 6695523; JOINED.
DR EMBL; X79478; 6695523; JOINED.
DR EMBL; X79479; 6695523; JOINED.
DR EMBL; X79480; 6695523; JOINED.
DR EMBL; X79481; 6695523; JOINED.
DR EMBL; X79482; 6695523; JOINED.
DR EMBL; X79483; 6695523; JOINED.
DR EMBL; X79484; 6695523; JOINED.
DR EMBL; X79485; 6695523; JOINED.
DR EMBL; X79486; 6695523; JOINED.
DR EMBL; X79487; 6695523; JOINED.
DR EMBL; X79488; 6695523; JOINED.
DR EMBL; X79489; 6695523; JOINED.
DR EMBL; X79490; 6695523; JOINED.
DR EMBL; X79491; 6695523; JOINED.
DR EMBL; X79492; 6695523; JOINED.
DR EMBL; X79493; 6695523; JOINED.
DR EMBL; X79494; 6695523; JOINED.
DR EMBL; X79495; 6695523; JOINED.
DR EMBL; X79496; 6695523; JOINED.
DR EMBL; X79497; 6695523; JOINED.
DR EMBL; X79498; 6695523; JOINED.
DR EMBL; X79499; 6695523; JOINED.
DR EMBL; X79500; 6695523; JOINED.
DR EMBL; X79501; 6695523; JOINED.
DR EMBL; X79502; 6695523; JOINED.
DR EMBL; X79503; 6695523; JOINED.
DR EMBL; X79504; 6695523; JOINED.
DR EMBL; X79505; 6695523; JOINED.
DR EMBL; X79506; 6695523; JOINED.
DR EMBL; X79507; 6695523; JOINED.
DR EMBL; X79508; 6695523; JOINED.
DR EMBL; X79509; 6695523; JOINED.
DR EMBL; X79510; 6695523; JOINED.
DR EMBL; X79511; 6695523; JOINED.
DR EMBL; X79512; 6695523; JOINED.
DR EMBL; X79513; 6695523; JOINED.
DR EMBL; X79514; 6695523; JOINED.
DR EMBL; X79515; 6695523; JOINED.
DR EMBL; X79516; 6695523; JOINED.
DR EMBL; X79517; 6695523; JOINED.
DR EMBL; X79518; 6695523; JOINED.
DR EMBL; X79519; 6695523; JOINED.
DR EMBL; X79520; 6695523; JOINED.
DR EMBL; X79521; 6695523; JOINED.
DR EMBL; X79522; 6695523; JOINED.
DR EMBL; X79523; 6695523; JOINED.
DR EMBL; X79524; 6695523; JOINED.
DR EMBL; X79525; 6695523; JOINED.
DR EMBL; X79526; 6695523; JOINED.
DR EMBL; X79527; 6695523; JOINED.
DR EMBL; X79528; 6695523; JOINED.
DR EMBL; X79529; 6695523; JOINED.
DR EMBL; X79530; 6695523; JOINED.
DR EMBL; X79531; 6695523; JOINED.
DR EMBL; X79532; 6695523; JOINED.
DR EMBL; X79533; 6695523; JOINED.
DR EMBL; X79534; 6695523; JOINED.
DR EMBL; X79535; 6695523; JOINED.
DR EMBL; X79536; 6695523; JOINED.
DR EMBL; X79537; 6695523; JOINED.
DR EMBL; X79538; 6695523; JOINED.
DR EMBL; X79539; 6695523; JOINED.
DR EMBL; X79540; 6695523; JOINED.
DR EMBL; X79541; 6695523; JOINED.
DR EMBL; X79542; 6695523; JOINED.
DR EMBL; X79543; 6695523; JOINED.
DR EMBL; X79544; 6695523; JOINED.
DR EMBL; X79545; 6695523; JOINED.
DR EMBL; X79546; 6695523; JOINED.
DR EMBL; X79547; 6695523; JOINED.
DR EMBL; X79548; 6695523; JOINED.
DR EMBL; X79549; 6695523; JOINED.
DR EMBL; X79550; 6695523; JOINED.
DR EMBL; X79551; 6695523; JOINED.
DR EMBL; X79552; 6695523; JOINED.
DR EMBL; X79553; 6695523; JOINED.
DR EMBL; X79554; 6695523; JOINED.
DR EMBL; X79555; 6695523; JOINED.
DR EMBL; X79556; 6695523; JOINED.
DR EMBL; X79557; 6695523; JOINED.
DR EMBL; X79558; 6695523; JOINED.
DR EMBL; X79559; 6695523; JOINED.
DR EMBL; X79560; 6695523; JOINED.
DR EMBL; X79561; 6695523; JOINED.
DR EMBL; X79562; 6695523; JOINED.
DR EMBL; X79563; 6695523; JOINED.
DR EMBL; X79564; 6695523; JOINED.
DR EMBL; X79565; 6695523; JOINED.
DR EMBL; X79566; 6695523; JOINED.
DR EMBL; X79567; 6695523; JOINED.
DR EMBL; X79568; 6695523; JOINED.
DR EMBL; X79569; 6695523; JOINED.
DR EMBL; X79570; 6695523; JOINED.
DR EMBL; X79571; 6695523; JOINED.
DR EMBL; X79572; 6695523; JOINED.
DR EMBL; X79573; 6695523; JOINED.
DR EMBL; X79574; 6695523; JOINED.
DR EMBL; X79575; 6695523; JOINED.
DR EMBL; X79576; 6695523; JOINED.
DR EMBL; X79577; 6695523; JOINED.
DR EMBL; X79578; 6695523; JOINED.
DR EMBL; X79579; 6695523; JOINED.
DR EMBL; X79580; 6695523; JOINED.
DR EMBL; X79581; 6695523; JOINED.
DR EMBL; X79582; 6695523; JOINED.
DR EMBL; X79583; 6695523; JOINED.
DR EMBL; X79584; 6695523; JOINED.
DR EMBL; X79585; 6695523; JOINED.
DR EMBL; X79586; 6695523; JOINED.
DR EMBL; X79587; 6695523; JOINED.
DR EMBL; X79588; 6695523; JOINED.
DR EMBL; X79589; 6695523; JOINED.
DR EMBL; X79590; 6695523; JOINED.
DR EMBL; X79591; 6695523; JOINED.
DR EMBL; X79592; 6695523; JOINED.
DR EMBL; X79593; 6695523; JOINED.
DR EMBL; X79594; 6695523; JOINED.
DR EMBL; X79595; 6695523; JOINED.
DR EMBL; X79596; 6695523; JOINED.
DR EMBL; X79597; 6695523; JOINED.
DR EMBL; X79598; 6695523; JOINED.
DR EMBL; X79599; 6695523; JOINED.
DR EMBL; X79600; 6695523; JOINED.
DR EMBL; X79601; 6695523; JOINED.
DR EMBL; X79602; 6695523; JOINED.
DR EMBL; X79603; 6695523; JOINED.
DR EMBL; X79604; 6695523; JOINED.
DR EMBL; X79605; 6695523; JOINED.
DR EMBL; X79606; 6695523; JOINED.
DR EMBL; X79607; 6695523; JOINED.
DR EMBL; X79608; 6695523; JOINED.
DR EMBL; X79609; 6695523; JOINED.
DR EMBL; X79610; 6695523; JOINED.
DR EMBL; X79611; 6695523; JOINED.
DR EMBL; X79612; 6695523; JOINED.
DR EMBL; X79613; 6695523; JOINED.
DR EMBL; X79614; 6695523; JOINED.
DR EMBL; X79615; 6695523; JOINED.
DR EMBL; X79616; 6695523; JOINED.
DR EMBL; X79617; 6695523; JOINED.
DR EMBL; X79618; 6695523; JOINED.
DR EMBL; X79619; 6695523; JOINED.
DR EMBL; X79620; 6695523; JOINED.
DR EMBL; X79621; 6695523; JOINED.
DR EMBL; X79622; 6695523; JOINED.
DR EMBL; X79623; 6695523; JOINED.
DR EMBL; X79624; 6695523; JOINED.
DR EMBL; X79625; 6695523; JOINED.
DR EMBL; X79626; 6695523; JOINED.
DR EMBL; X79627; 6695523; JOINED.
DR EMBL; X79628; 6695523; JOINED.
DR EMBL; X79629; 6695523; JOINED.
DR EMBL; X79630; 6695523; JOINED.
DR EMBL; X79631; 6695523; JOINED.
DR EMBL; X79632; 6695523; JOINED.
DR EMBL; X79633; 6695523; JOINED.
DR EMBL; X79634; 6695523; JOINED.
DR EMBL; X79635; 6695523; JOINED.
DR EMBL; X79636; 6695523; JOINED.
DR EMBL; X79637; 6695523; JOINED.
DR EMBL; X79638; 6695523; JOINED.
DR EMBL; X79639; 6695523; JOINED.
DR EMBL; X79640; 6695523; JOINED.
DR EMBL; X79641; 6695523; JOINED.
DR EMBL; X79642; 6695523; JOINED.
DR EMBL; X79643; 6695523; JOINED.
DR EMBL; X79644; 6695523; JOINED.
DR EMBL; X79645; 6695523; JOINED.
DR EMBL; X79646; 6695523; JOINED.
DR EMBL; X79647; 6695523; JOINED.
DR EMBL; X79648; 6695523; JOINED.
DR EMBL; X79649; 6695523; JOINED.
DR EMBL; X79650; 6695523; JOINED.
DR EMBL; X79651; 6695523; JOINED.
DR EMBL; X79652; 6695523; JOINED.
DR EMBL; X79653; 6695523; JOINED.
DR EMBL; X79654; 6695523; JOINED.
DR EMBL; X79655; 6695523; JOINED.
DR EMBL; X79656; 6695523; JOINED.
DR EMBL; X79657; 6695523; JOINED.
DR EMBL; X79658; 6695523; JOINED.
DR EMBL; X79659; 6695523; JOINED.
DR EMBL; X79660; 6695523; JOINED.
DR EMBL; X79661; 6695523; JOINED.
DR EMBL; X79662; 6695523; JOINED.
DR EMBL; X79663; 6695523; JOINED.
DR EMBL; X79664; 6695523; JOINED.
DR EMBL; X79665; 6695523; JOINED.
DR EMBL; X79666; 6695523; JOINED.
DR EMBL; X79667; 6695523; JOINED.
DR EMBL; X79668; 6695523; JOINED.
DR EMBL; X79669; 6695523; JOINED.
DR EMBL; X79670; 6695523; JOINED.
DR EMBL; X79671; 6695523; JOINED.
DR EMBL; X79672; 6695523; JOINED.
DR EMBL; X79673; 6695523; JOINED.
DR EMBL; X79674; 6695523; JOINED.
DR EMBL; X79675; 6695523; JOINED.
DR EMBL; X79676; 6695523; JOINED.
DR EMBL; X79677; 6695523; JOINED.
DR EMBL; X79678; 6695523; JOINED.
DR EMBL; X79679; 6695523; JOINED.
DR EMBL; X79680; 6695523; JOINED.
DR EMBL; X79681; 6695523; JOINED.
DR EMBL; X79682; 6695523; JOINED.
DR EMBL; X79683; 6695523; JOINED.
DR EMBL; X79684; 6695523; JOINED.
DR EMBL; X79685; 6695523; JOINED.
DR EMBL; X79686; 6695523; JOINED.
DR EMBL; X79687; 6695523; JOINED.
DR EMBL; X79688; 6695523; JOINED.
DR EMBL; X79689; 6695523; JOINED.
DR EMBL; X79690; 6695523; JOINED.
DR EMBL; X79691; 6695523; JOINED.
DR EMBL; X79692; 6695523; JOINED.
DR EMBL; X79693; 6695523; JOINED.
DR EMBL; X79694; 6695523; JOINED.
DR EMBL; X79695; 6695523; JOINED.
DR EMBL; X79696; 6695523; JOINED.
DR EMBL; X79697; 6695523; JOINED.
DR EMBL; X79698; 6695523; JOINED.
DR EMBL; X79699; 6695523; JOINED.
DR EMBL; X79700; 6695523; JOINED.
DR EMBL; X79701; 6695523; JOINED.
DR EMBL; X79702; 6695523; JOINED.
DR EMBL; X79703; 6695523; JOINED.
DR EMBL; X79704; 6695523; JOINED.
DR EMBL; X79705; 6695523; JOINED.
DR EMBL; X79706; 6695523; JOINED.
DR EMBL; X79707; 6695523; JOINED.
DR EMBL; X79708; 6695523; JOINED.
DR EMBL; X79709; 6695523; JOINED.
DR EMBL; X79710; 6695523; JOINED.
DR EMBL; X79711; 6695523; JOINED.
DR EMBL; X79712; 6695523; JOINED.
DR EMBL; X79713; 6695523; JOINED.
DR EMBL; X79714; 6695523; JOINED.
DR EMBL; X79715; 6695523; JOINED.
DR EMBL; X79716; 6695523; JOINED.
DR EMBL; X79717; 6695523; JOINED.
DR EMBL; X79718; 6695523; JOINED.
DR EMBL; X79719; 6695523; JOINED.
DR EMBL; X79720; 6695523; JOINED.
DR EMBL; X79721; 6695523; JOINED.
DR EMBL; X79722; 6695523; JOINED.
DR EMBL; X79723; 6695523; JOINED.
DR EMBL; X79724; 6695523; JOINED.
DR EMBL; X79725; 6695523; JOINED.
DR EMBL; X79726; 6695523; JOINED.
DR EMBL; X79727; 6695523; JOINED.
DR EMBL; X79728; 6695523; JOINED.
DR EMBL; X79729; 6695523; JOINED.
DR EMBL; X79730; 6695523; JOINED.
DR EMBL; X79731; 6695523; JOINED.
DR EMBL; X79732; 6695523; JOINED.
DR EMBL; X79733; 6695523; JOINED.
DR EMBL; X79734; 6695523; JOINED.
DR EMBL; X79735; 6695523; JOINED.
DR EMBL; X79736; 6695523; JOINED.
DR EMBL; X79737; 6695523; JOINED.
DR EMBL; X79738; 6695523; JOINED.
DR EMBL; X79739; 6695523; JOINED.
DR EMBL; X79740; 6695523; JOINED.
DR EMBL; X79741; 6695523; JOINED.
DR EMBL; X79742; 6695523; JOINED.
DR EMBL; X79743; 6695523; JOINED.
DR EMBL; X79744; 6695523; JOINED.
DR EMBL; X79745; 6695523; JOINED.
DR EMBL; X79746; 6695523; JOINED.
DR EMBL; X79747; 6695523; JOINED.
DR EMBL; X79748; 6695523; JOINED.
DR EMBL; X79749; 6695523; JOINED.
DR EMBL; X79750; 6695523; JOINED.
DR EMBL; X79751; 6695523; JOINED.
DR EMBL; X79752; 6695523; JOINED.
DR EMBL; X79753; 6695523; JOINED.
DR EMBL; X79754; 6695523; JOINED.
DR EMBL; X79755; 6695523; JOINED.
DR EMBL; X79756; 6695523; JOINED.
DR EMBL; X79757; 6695523; JOINED.
DR EMBL; X79758; 6695523; JOINED.
DR EMBL; X79759; 6695523; JOINED.
DR EMBL; X79760; 6695523; JOINED.
DR EMBL; X79761; 6695523; JOINED.
DR EMBL; X79762; 6695523; JOINED.
DR EMBL; X79763; 6695523; JOINED.
DR EMBL; X79764; 6695523; JOINED.
DR EMBL; X79765; 6695523; JOINED.
DR EMBL; X79766; 6695523; JOINED.
DR EMBL; X79767; 6695523; JOINED.
DR EMBL; X79768; 6695523; JOINED.
DR EMBL; X79769; 6695523; JOINED.
DR EMBL; X79770; 6695523; JOINED.
DR EMBL; X79771; 6695523

(a) (b) (c) (d) (e) (f)

CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC ORIGINAL SOURCE:
CC ORGANISM: human
SQ SEQUENCE 269 AA; 30227 MW; 341084 CN;

Query Match
Best Local Similarity 81.8%; Score 54; DB 1; Length 269;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 127 ORGNRPY 134
111111
QY 1 ORGNRPY 8

RESULT 2
ID 5198342-2 STANDARD; PRT; 311 AA.
XX xxxxxx
DT 01-JAN-1900
XX Patent No. 5198342.
DE Patent No. 5198342.
CC Patent No. 5198342
CC APPLICANT: MALISZEWSKI, CHARLES R.
CC TITLE OF INVENTION: DNA ENCODING IGA FC RECEPTORS
CC NUMBER OF SEQUENCES: 9
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/548,059
CC FILING DATE: 05-JUL-1990
CC SEQ ID NO: 2:
DE LENGTH: 287
SQ SEQUENCE 311 AA; 34908 MW; 558002 CN;

Query Match
Best Local Similarity 75.8%; Score 50; DB 4; Length 287;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 196 WYNRSFY 202
:111111
QY 2 RYNRPY 8

RESULT 3
ID US-07-971-092-2 STANDARD; PRT; 287 AA.
XX xxxxxx
DT
XX
DE Sequence 2, Application US/07971092
XX
CC Sequence 2, Application US/07971092
CC Patent No. 5328987
CC GENERAL INFORMATION:
CC APPLICANT: Maliszewski, Charles R.
CC TITLE OF INVENTION: Huiga Fc Receptor
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunex
CC STREET: 51 University
CC CITY: Seattle
CC STATE: WA
CC COUNTRY: USA
CC ZIP: 98101
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/07/971,092
CC FILING DATE: 19921104
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Perkins, Patricia A.
CC REGISTRATION NUMBER: 34693
CC REFERENCE/DOCKET NUMBER: 2603
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 287 amino acids
CC TYPE: AMINO ACID
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 287 AA; 32265 MW; 435416 CN;

Query Match
Best Local Similarity 75.8%; Score 50; DB 1; Length 287;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 196 WYNRSFY 202
:111111
QY 2 RYNRPY 8

RESULT 4
ID US-08-604-989A-3 STANDARD; PRT; 246 AA.
XX xxxxxx
AC
XX
DT
XX
DE Sequence 3, Application US/08604989A
XX
CC Patent No. 5834208
CC GENERAL INFORMATION:
CC APPLICANT: Sakano, S.
CC TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
CC NUMBER OF SEQUENCES: 11
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds LLP
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/604,989A
CC FILING DATE: February 23, 1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Charles E. Miller
CC REGISTRATION NUMBER: 24,576
CC REFERENCE/DOCKET NUMBER: 1920-026
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-8864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 246 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC ORIGINAL SOURCE:
CC ORGANISM: human
CC STRAIN: UT-7
SQ SEQUENCE 246 AA; 27425 MW; 326654 CN;

Query Match 68.2%; Score 45; DB 2; Length 246;
Best Local Similarity 83.3%; Pred. No. 1.30e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 188 YGRAPY 193
1 ||||
3 YNRAPY 8

RESULT 5
ID US-08-604-989A-4 STANDARD; PRT; 466 AA.

Sequence 4, Application US/08604989A

Sequence 4, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:

APPLICANT: Sakano, S.
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7

SEQUENCE 466 AA; 51898 MW; 1114727 CN;
Query Match 68.2%; Score 45; DB 2; Length 466;
Best Local Similarity 83.3%; Pred. No. 1.30e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 379 YGRAPY 384
1 ||||
3 YNRAPY 8

RESULT 6
ID PCT-US95-05008-2 STANDARD; PRT; 507 AA.
AC xxxxxx

Sequence 2, Application PC/TUS9505008

GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Hoigarten Str. 2
APPLICANT: Munchen 80539
APPLICANT: Germany
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-074
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE 507 AA; 56469 MW; 1316735 CN;

Query Match 68.2%; Score 45; DB 3; Length 507;
Best Local Similarity 83.3%; Pred. No. 1.30e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 420 YGRAPY 425
1 ||||
3 YNRAPY 8

RESULT 7
ID US-08-604-989A-5 STANDARD; PRT; 507 AA.
AC xxxxxx

Sequence 5, Application US/08604989A

CC Sequence 5, Application US/08604989A
CC Patent No. 5834208
CC GENERAL INFORMATION:
CC APPLICANT: Sakano, S.
CC TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
CC NUMBER OF SEQUENCES: 11
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennile & Edmonds LLP
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/604 989A
CC FILING DATE: February 23, 1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Charles E. Miller
CC REGISTRATION NUMBER: 24,576
CC REFERENCE/DOCKET NUMBER: 1920-026
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-8864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 507 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC ORIGINAL SOURCE:
CC ORGANISM: human
CC STRAIN: UT-7
CC QO SEQUENCE 507 AA; 56491 MW; 1317560 CN;

Query Match 68.2%; Score 45; DB 2; Length 507;
Best Local Similarity 83.3%; Pred. No. 130e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 420 YGRAPY 425
| | | |
3 YNRARY 8

RESULT 8
ID US-08-724-194-12 STANDARD: PRT: 33 AA.
AC xxxxxx
DT
XX
DE Sequence 12, Application US/08724194
XX
CC Sequence 12, Application US/08724194
CC Patent No. 5824875
CC GENERAL INFORMATION:
CC APPLICANT: RANU, RAJINDER S.
CC TITLE OF INVENTION: ONE-AMINOACYCLOPROPANE-1-CARBOXYLATE
CC TITLE OF INVENTION: SYNTHASE GENES FROM PELARGONIUM TO CONTROL ETHYLENE LEVELS
CC TITLE OF INVENTION: IN GERANIUMS
CC NUMBER OF SEQUENCES: 13
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SANTANGELO LAW OFFICES PC
CC STREET: 315 WEST OAK STREET, STE 701
CC CITY: FORT COLLINS
CC STATE: CO
CC COUNTRY: USA

N/L

```

CC      ZIP: 80521
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/724,194
CC      FILING DATE: 01-OCT-1996
CC      CLASSIFICATION: 800
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: SANTANGELO, LUKE
CC      REGISTRATION NUMBER: 31,997
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (970) 224-3100
CC      INFORMATION FOR SEQ ID NO: 12:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 33 amino acids
CC      TYPE: amino acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: peptide
CC      SEQUENCE 33 AA: 3822 MW; 5035 CN;
SQ
Query Match          66.7%; Score 44; DB 2; Length 33;
Best Local Similarity 66.7%; Pred. No. 1.65e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db      8 YDRDPY 13
QY      3 YNRAPY 8
RESULT 9
ID      US-08-343-101A-20 STANDARD; PRT; 218 AA.
XX      xxxxxx
XX
DE      Sequence 20, Application US/08343101A
AC      Sequence 20, Application US/08343101A
AC      Patent No. 5830759
CC      GENERAL INFORMATION:
CC      APPLICANT: Chang, Yuan
CC      APPLICANT: Moore, Patrick S.
CC      TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma
CC      TITLE OF INVENTION: Virus Sequences And Uses Thereof
CC      NUMBER OF SEQUENCES: 22
CC      CORRESPONDENCE ADDRESSES:
CC      ADDRESSEE: Cooper & Dunham
CC      STREET: 1185 Avenue of the Americas
CC      CITY: New York
CC      STATE: New York
CC      COUNTRY: U.S.A.
CC      ZIP: 10036
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.24
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/343,101A
CC      FILING DATE:
CC      CLASSIFICATION: 514
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: White Esq., John P.
CC      REGISTRATION NUMBER: 28,678
CC      REFERENCE/DOCKET NUMBER: 45185-A
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 212-278-0400
CC      TELEFAX: 212-391-0526
CC

```

CC INFORMATION FOR SEQ ID NO: 20:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 218 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC HYPOTHETICAL: N
CC ANTI-SENSE: N
CC FEATURE:
CC NAME/KEY: Peptide
CC LOCATION: 1..218
CC OTHER INFORMATION:
SO SEQUENCE 218 AA; 24345 MW; 253377 CN;

Query Match 66.7%; Score 44; DB 2; Length 218;
Best Local Similarity 71.4%; Pred. No. 1.65e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

119 RYNRGLY 125
||||:|
2 RYNRAPY 8

RESULT 10
ID US-07-989-991A-2 STANDARD: PRT; 400 AA.
XX
AC xxxxxx
XX
DT
DE
XX
XX
Sequence 2, Application US/07989991A
CC
CC Sequence 2, Application US/07989991A
CC Patent No. 5525502
CC GENERAL INFORMATION:
CC APPLICANT: Thireos, George
CC APPLICANT: Kafetzopoulos, Dimitris
CC TITLE OF INVENTION: DNA ENCODING CHITIN DEACETYLASE
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Militia Drive
CC CITY: Lexington
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/989,991A
CC FILING DATE: 07-DEC-1992
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Brook, David E.
CC REGISTRATION NUMBER: 22,552
CC REFERENCE/DOCKET NUMBER: BT91-01A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617 861-6240
CC TELEFAX: 617 861-9540
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 400 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SO SEQUENCE 400 AA; 43857 MW; 860821 CN;

Query Match 66.7%; Score 44; DB 1; Length 400;
Best Local Similarity 71.4%; Pred. No. 1.65e+02;

NR

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 228 RYNRPPY 234
||||:
OY 2 RYNRAPY 8

RESULT 11
ID US-08-748-485-1 STANDARD: PRT; 454 AA.
XX
AC xxxxxx
XX
DT
DE
XX
XX
Sequence 1, Application US/08748485
CC
CC Sequence 1, Application US/08748485
CC Patent No. 5817480
CC GENERAL INFORMATION:
CC APPLICANT: Au-Young, Janice
CC APPLICANT: Guegler, Karl J.
CC APPLICANT: Goli, Surya K.
CC APPLICANT: Murty, Lynn E.
CC TITLE OF INVENTION: NOVEL HISTAMINE H2 RECEPTOR
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: US
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/748,485
CC FILING DATE: Herewith
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: PF-0159 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 454 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: Consensus
CC CLONE: 1722180
SO SEQUENCE 454 AA; 50589 MW; 1146364 CN;

Query Match 66.7%; Score 44; DB 2; Length 454;
Best Local Similarity 50.0%; Pred. No. 1.65e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 343 DRYRREP 350
:|||:
OY 1 ORYNAPY 8

RESULT 12

NR

ID US-08-467-568-2 STANDARD; PRT: 529 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 2, Application US/08467568
XX
CC Sequence 2, Application US/08467568
CC Patent No. 5817477
CC GENERAL INFORMATION:
CC APPLICANT: SOPPET, DANIEL R
CC TITLE OF INVENTION: ADRENALGIC RECEPTOR
CC NUMBER OF SEQUENCES: 13
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
CC ADDRESSEE: Stewart & Olstein
CC STREET: 6 Becker Farm Road
CC CITY: Roseland
CC STATE: NJ
CC COUNTRY: USA
CC ZIP: 07068-1739
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467, 568
CC FILING DATE: 06-JUN-1995
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Ferraro, Gregory D
CC REGISTRATION NUMBER: 36,134
CC REFERENCE/DOCKET NUMBER: 325800-324
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 201-994-1700
CC TELEFAX: 201-994-1744
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 529 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: Protein
CC SEQUENCE 529 AA; 58358 MW; 1481658 CN;
SQ
Query Match 66.7%; Score 44; DB 2; Length 529;
Best Local Similarity 50.0%; Pred. No. 1.65e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
DB 342 DRYREPF 349
QY 1 QRYNRPY 8
RESULT 13
ID PCT-US94-09051-2 STANDARD; PRT: 529 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 2, Application PC/TUS9409051
XX
CC Sequence 2, Application PC/TUS9409051
CC GENERAL INFORMATION:
CC APPLICANT: LI, ET AL.
CC TITLE OF INVENTION: Adrenergic Receptor
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
CC ADDRESSEE: CECCHI, STEWART & OLSTEIN
CC STREET: 6 BECKER FARM ROAD
CC

CC CITY: ROSELAND
CC STATE: NEW JERSEY
CC COUNTRY: USA
CC ZIP: 07068
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3.5 INCH DISKETTE
CC COMPUTER: IBM PS/2
CC OPERATING SYSTEM: MS-DOS
CC SOFTWARE: WORD PERFECT 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/09051
CC FILING DATE: Submitted herewith
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: FERRARO, GREGORY D
CC REGISTRATION NUMBER: 36,134
CC REFERENCE/DOCKET NUMBER: 325800-194
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 201-994-1700
CC TELEFAX: 201-994-1744
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 529 AMINO ACIDS
CC TYPE: AMINO ACID
CC STRANDEDNESS:
CC TOPOLOGY: LINEAR
CC MOLECULE TYPE: PROTEIN
CC SEQUENCE 529 AA; 58358 MW; 1481658 CN;
SQ
Query Match 66.7%; Score 44; DB 3; Length 529;
Best Local Similarity 50.0%; Pred. No. 1.65e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
DB 342 DRYREPF 349
QY 1 QRYNRPY 8
RESULT 14
ID PCT-US95-10194-3 STANDARD; PRT: 1376 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 3, Application PC/TUS9510194
XX
CC Sequence 3, Application PC/TUS9510194
CC GENERAL INFORMATION:
CC APPLICANT: The Trustees of Columbia University in the City of New York
CC APPLICANT: City
CC TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
CC NUMBER OF SEQUENCES: 45
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Cooper & Dunham LLP
CC STREET: 1185 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/10194
CC FILING DATE:
CC CLASSIFICATION:

CC ATTORNEY/AGENT INFORMATION:
CC NAME: White, John P.
CC REGISTRATION NUMBER: 28, 678
CC REFERENCE/DOCKET NUMBER: 45185-C-PCT/JPM/NSC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 278-0400
CC TELEFAX: (212) 391-0525
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1376 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1376 AA; 153401 MW; 9818578 CN;

Query Match 66.7%; Score 44; DB 3; Length 1376;
Best Local Similarity 71.4%; Pred. No. 1.65e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1277 RYNRGLY 1283
||||:|
2 RYNRAPY 8

RESULT 15
ID US-08-420-235B-3 STANDARD; PRT; 1376 AA.
XX xxxxxx
DT
XX
DE
XX
XX
Sequence 3, Application US/08420235B

Sequence 3, Application US/08420235B
Patent No. 5801042

GENERAL INFORMATION:

APPLICANT: Chang, Yuan

APPLICANT: Moore, Patrick S.

TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS

TITLE OF INVENTION: SEQUENCES AND USES THEREOF

NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/420,235B

FILING DATE:

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28, 678

REFERENCE/DOCKET NUMBER: 45185-B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1376 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 1376 AA; 153401 MW; 9818578 CN;

Query Match 66.7%; Score 44; DB 2; Length 1376;

Best Local Similarity 71.4%; Pred. No. 1.65e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 1277 RYNRGLY 1283
||||:|
QY 2 RYNRAPY 8

Search completed: Thu Sep 2 11:19:10 1999
Job time : 8 secs.

THIS PAGE BLANK (USPTO)

 WISE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit,
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Run on: Thu Sep 2 11:17:30 1999; Maspar time 3.08 Seconds
 116.921 Million cell updates/sec
 MPerch_PP protein - protein database search, using Smith-Waterman algorithm
 Molecular output not generated.

Title: >US-08-599-226-3
 Description: (1-9) from US08599226..pep
 Perfect Score: 66
 Sequence: 1 ORYRAPPYX 9

Scoring table: PAM 150
 Gap 15

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: p1r60
 1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 22.973; Variance 30.279; scale 0.759

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	57	86.4	279	1	SNF5K	multicatalytic endope	3.45e+01
2	54	81.8	263	1	SNR1C2	multicatalytic endope	1.43e+00
3	54	81.8	269	2	UC1445	multicatalytic endope	1.43e+00
4	53	80.3	790	2	S77032	ABC-type transport pr	2.27e+00
5	52	78.8	321	2	B71854	type II DNA modifi	3.58e+00
6	51	77.3	1260	2	S28407	guanine nucleotide-ex	5.63e+00
7	50	75.8	239	2	G02630	Fcalphab - human	8.80e+00
8	50	75.8	287	2	JH0332	IGA (Fc) receptor, my	8.80e+00
9	50	75.8	521	2	I39956	neutral proteinase (E	8.80e+00
10	49	74.2	350	2	S52662	S-adenosylmethionine	1.37e+01
11	49	74.2	360	2	S28047	TUBB3 protein - potat	1.37e+01
12	48	72.7	367	2	S49009	fork head protein 2 -	2.11e+01
13	48	72.7	370	2	S49008	fork head protein - A	2.11e+01
14	48	72.7	534	2	S71800	transcription factor	2.11e+01
15	48	72.7	1032	2	S74487	hypothetical protein	3.24e+01
16	47	71.2	14	1	NYPG14	hypothetical tetrade	3.24e+01
17	47	71.2	288	2	S22628	hypothetical protein	3.24e+01
18	47	71.2	323	2	C56024	GDP dissociation inh	3.24e+01
19	47	71.2	355	2	S76940	hypothetical protein	3.24e+01
20	47	71.2	440	2	F70792	hypothetical protein	3.24e+01
21	47	71.2	442	1	VJFF2	vitellogenin II precu	3.24e+01
22	47	71.2	445	2	C56956	GDP dissociation inh	3.24e+01
23	47	71.2	445	2	B54091	rab GDP dissociation	3.24e+01

24	47	71.2	445	2	A56024	GDP dissociation inh	3.24e+01
25	47	71.2	447	2	A54091	rab GDP dissociation	3.24e+01
26	47	71.2	447	2	A35652	smg p25a regulatory p	3.24e+01
27	47	71.2	447	2	I37082	GDP-dissociation inh	3.24e+01
28	47	71.2	447	2	B56024	GDP dissociation inh	3.24e+01
29	47	71.2	448	2	S36746	GDP dissociation inh	3.24e+01
30	47	71.2	451	2	S44446	GDP dissociation inh	3.24e+01
31	47	71.2	515	2	T03070	hypothetical protein	3.24e+01
32	47	71.2	520	2	B71143	hypothetical protein	3.24e+01
33	47	71.2	541	2	S01957	hypothetical protein	3.24e+01
34	47	71.2	689	2	S45901	probable membrane pro	3.24e+01
35	47	71.2	722	2	C71411	hypothetical protein	3.24e+01
36	47	71.2	1320	2	S66279	proline dehydrogenase	3.24e+01
37	47	71.2	1320	2	D64843	proline dehydrogenase	3.24e+01
38	46	69.7	106	2	G02071	helicase - human (f	4.93e+01
39	46	69.7	182	2	S45202	hypoxanthine phosphor	4.93e+01
40	46	69.7	385	2	C37753	quinoxaline synthase	4.93e+01
41	46	69.7	419	3	JE0389	catalbrite repressor	4.93e+01
42	46	69.7	613	2	S55615	thymidine kinase (EC	4.93e+01
43	46	69.7	1409	2	S41028	hypothetical protein	4.93e+01
44	45	68.2	527	2	A49865	protein-tyrosine kin	7.47e+01
45	45	68.2	544	2	S41093	triacylglycerol lipas	7.47e+01

ALIGNMENTS

RESULT	1	ALIGNMENTS
ENTRY	SNF5K	#type complete
TITLE	multicatalytic endopeptidase complex (EC 3.4.99.46) 35k chain	
ALTERNATE_NAMES	- fruit fly (Drosophila melanogaster)	
ORGANISM	19S cylinder particle 35k chain: multicatalytic proteinase	
DATE	35k chain: prosome 35k chain: proteasome 35k chain	
ACCESSIONS	#formal name Drosophila melanogaster	
REFERENCE	30-Sep-1991 #sequence_revision 30-Sep-1991 #text-change	
#authors	05-Sep-1997	
#journal	S23450; S05507; A38761	
#title	S23450	
#cross-references	EMBL:X62285; NID:g8387; PID:g8388	
#molecule_type	DNA	
#residues	1-279 ##label FREN	
#cross-references	EMBL:X62285; NID:g8387; PID:g8388	
#experimental_source	strain Canton S	
#accession	S05507	
#molecule_type	RNA	
#residues	1-279 ##label HAA	
#cross-references	EMBL:X15497; NID:g8381; PID:g8382	
#accession	A38761	
#molecule_type	protein	
#residues	4-18;194-206 ##label HAA2	
#gene	PROS-35	
#cross-references	flybase:FBgn0003151	
#map position	89F-90A	
#intron	1/3; 211/3	
CLASSIFICATION	#superfamily multicatalytic endopeptidase complex chain C9	
KEYWORDS	hydrolase; phosphoprotein; proteinase	
SUMMARY	#length 279 #molecular-weight 31058 #checksum 365	
Query Match	86.4%; Score 57; DB 1; Length 279;	

Best Local Similarity 75.0%; Pred. No. 3.45e-01;
Matches 6; Conservatave 1; Mismatches 1; Indels 0; Gaps 0;

Db 121 ORYDRAPY 128
11111111
OY 1 ORYMRAPY 8

RESULT 2
ENTRY SNRTC2 #type complete
TITLE multicatalytic endopeptidase complex (EC 3.4.99.46) chain C2

ALTERNATE_NAMES multicatalytic proteinase component C2; proteasome chain C2
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Oct-1997

ACCESSIONS A32968; A38799; S09741
REFERENCE A32968
#authors Fujiwara, T.; Tanaka, K.; Kumatori, A.; Shin, S.; Yoshimura, T.; Ichihara, A.; Tokunaga, F.; Aruga, R.; Iwanaga, S.; Kakiyuka, A.; Nakajishi, S.
#journal Biochemistry (1989) 28:7332-7340
#title Molecular cloning of cDNA for proteasomes (multicatalytic proteinase complexes) from rat liver: primary structure of the largest component (C2).

#cross-references MIMD:90057428
#accession A32968
#molecule_type mRNA
#residues 1-263 ##label FUJ1
#cross-references EMBL:M29859; NID:g206381; PID:g206382
#accession A38799
#molecule_type protein
#residues 2-25;42-58;63-74,'X',76-79,'X',81,116-135;190-203;218-226,'XX',229,'X',231;244-246,'X',248-262 ##label FUJ2

REFERENCE S09741
#authors Tokunaga, F.; Aruga, R.; Iwanaga, S.; Tanaka, K.; Ichihara, A.; Takao, T.; Shimomishi, Y.
#journal FEBS Lett. (1990) 263:373-375
#title The NH2-terminal residues of rat liver proteasome (multicatalytic proteinase complex) subunits, C2, C3 and C8, are N-alpha-acetylated.
#cross-references MIMD:90243011
#accession S09741
#molecule_type protein
#residues 1-30 ##label TOK
CLASSIFICATION #superfamily multicatalytic endopeptidase complex chain C9
KEYWORDS acetylated amino end; hydrolase; proteinase
FEATURE 1-263 #product multicatalytic endopeptidase complex chain C2
#status experimental #label MAT
#modified_site acetylated amino end (Met) #status experimental

SUMMARY #length 263 #molecular-weight 29517 #checksum 9238

Query Match 81.8%; Score 54; DB 1; Length 263;
Best Local Similarity 75.0%; Pred. No. 1.43e+00;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 121 ORYDRAPY 128
11111111
OY 1 ORYMRAPY 8

RESULT 3
ENTRY JCI1445 #type complete
TITLE multicatalytic endopeptidase complex (EC 3.4.99.46) chain C2,
long splice form - human
ALTERNATE_NAMES macropain nu chain; multicatalytic endopeptidase complex HC2
chain; multicatalytic endopeptidase complex nu chain;
multicatalytic proteinase chain C2; prosome 30-33K chain;
prosome alpha 1 subunit; prosome chain C2; proteasome
nu chain

CONTAINS multicatalytic endopeptidase complex chain C2, short splice form

ORGANISM #formal_name Homo sapiens #common_name man
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Sep-1997

ACCESSIONS JCI1445; S15897; S17520; S25410; PC2321
REFERENCE JCI1445
#authors Pereira, I.S.; Bey, F.; Coux, O.; Scherrer, K.
#journal Gene (1993) 120:235-242
#title Two mRNAs exist for the Hs PROS-30 gene encoding a component of human prosome.

#accession JCI1445
#molecule_type mRNA
#residues 1-269 ##label PER
#cross-references GB:M64992
REFERENCE S15897
#authors Tamura, T.; Lee, D.H.; Osaka, F.; Fujiwara, T.; Shin, S.; Chung, C.H.; Tanaka, K.; Ichihara, A.
#journal Biochim. Biophys. Acta (1991) 1089:95-102
#title Molecular cloning and sequence analysis of cDNAs for five major subunits of human proteasomes (multi-catalytic proteinase complexes).

#cross-references MIMD:91223105
#accession S15897
#molecule_type mRNA
#residues 'M',8-269 ##label TAM
#cross-references EMBL:D00759; NID:g220021; PID:d100114; PID:g220022
REFERENCE S17520
#authors DeMartino, G.N.; Orth, K.; McCullough, M.L.; Lee, L.W.; Munn, T.Z.; Mooney, C.R.; Dawson, P.A.; Slaughter, C.A.
#journal Biochim. Biophys. Acta (1991) 1079:29-38
#title The primary structures of four subunits of the human, high-molecular-weight proteinase, macropain (prosome), are distinct but homologous.

#cross-references MIMD:91363412
#accession S17520
#status not compared with conceptual translation
#molecule_type mRNA
#residues 'M',8-269 ##label DEM
#cross-references GB:X61969; NID:g296737; PID:g296738
#accession S25410
#molecule_type protein
#residues 10-40;46-61;68-75;89-95;103-128;132-148;164-168;176-195;203-223;225-267 ##label DE2

REFERENCE PC2315
#authors Kristensen, P.; Johnsen, A.H.; Uerkytitz, W.; Tanaka, K.; Hendil, K.B.
#journal Biochem. Biophys. Res. Commun. (1994) 205:1785-1789
#title Human proteasome subunits from 2-dimensional gels identified by partial sequencing.
#cross-references MIMD:95110324
#accession PC2321
#molecule_type protein
#residues 69-88 ##label KRI
COMMENT #experimental_source placenta
The proteasome consists of subunits of 21K-30K arranged in 4 stacked rings.

GENETICS
#gene GDB:PSMAL1; HSPROS-30
#cross-references GDB:134040
CLASSIFICATION #superfamily multicatalytic endopeptidase complex chain C9
KEYWORDS hydrolase; phosphoprotein; proteinase
FEATURE 2-269 #product multicatalytic endopeptidase complex chain C2,
long splice form #status predicted #label MATV
KEYWORDS #product multicatalytic endopeptidase complex chain C2,
short splice form #status predicted #label MATS
230 #binding_site phosphate (Tyr) (covalent) #status predicted

SUMMARY #length 269 #molecular-weight 30239 #checksum 5946

Query Match 81.8%; Score 54; DB 2; Length 269;

Best Local Similarity 75.0%; Pred. No. 1.43e+00;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 127 QRYGRAPY 134
||| |
QY 1 QRYNRAPY 8

RESULT 4
ENTRY 577032 #type complete
TITLE ABC-type transport protein sl10778 - *Synechocystis* sp.
(strain PCC 6803)
ALTERNATE_NAMES protein sl10778
ORGANISM #formal_name *Synechocystis* sp.
#variety PCC 6803
DATE 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 21-Aug-1998

ACCESSIONS
REFERENCE S77032
#authors S74322
Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; Hirosewa, M.; Sugita, M.; Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.; Tabata, S.
#journal DNA Res. (1996) 3:109-136
#title Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.

#cross-references M01D:97061201
#accession S77032
#status nucleic acid sequence not shown; translation not shown
##residues 1-790 ##label KAN
##cross-references EMBL:D64005; GB:AB001339; NID:g1001779; PID:d1011375; PID:g1006575
#note the nucleotide sequence was submitted to the EMBL Data Library, June 1996

CLASSIFICATION #superfamily ATP-binding cassette homology
KEYWORDS P-loop: transport protein
FEATURE 244-434 #domain ATP-binding cassette homology #label ABC
261-268 #region nucleotide-binding motif A (P-loop)
SUMMARY #length 790 #molecular-weight 87656 #checksum 7114

Query Match 80.3%; Score 53; DB 2; Length 790;
Best Local Similarity 62.3%; Pred. No. 2.27e+00;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 475 QRYSRSSY 482
||| |
QY 1 QRYNRAPY 8

RESULT 5
ENTRY B71854 #type complete
TITLE type II DNA modification enzyme (methyltransferase) - *Helicobacter pylori* (strain J99)
ORGANISM #formal_name *Helicobacter pylori*
#variety strain J99
DATE 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 12-Feb-1999

ACCESSIONS
REFERENCE B71800
#authors B71854
Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.R.; Noonan, B.; Guild, B.C.; deJonge, B.L.; Carmel, G.; Tummino, P.J.; Carnuso, A.; Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Voyts, G.F.; Trust, T.J.
#journal Nature (1999) 397:176-180
#title Genomic sequence comparison of two unrelated isolates of the

human gastric pathogen *Helicobacter pylori*.
#cross-references M01D:99120557
#accession B71854
#status Preliminary
##molecule_type DNA
##residues 1-321 ##label ARN
##cross-references GB:AE001533; GB:AE001439; NID:g4155636; PID:g4155654
##experimental_source strain J99

GENETICS
#gene jhp1050
SUMMARY #length 321 #molecular-weight 36919 #checksum 5430

Query Match 78.8%; Score 52; DB 2; Length 321;
Best Local Similarity 75.0%; Pred. No. 3.58e+00;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 204 QRYLRNPY 211
||| |
QY 1 QRYNRAPY 8

RESULT 6
ENTRY S28407 #type complete
TITLE guanine nucleotide-exchange activator CDC25 homolog - mouse
ORGANISM #formal_name *Mus musculus* #common_name house mouse
DATE 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 10-Sep-1997

ACCESSIONS
REFERENCE S28407; S22693; B46199; S20730
#authors S28407
#journal Cen, H.; Papageorge, A.G.; Zippel, R.; Lowy, D.R.; Zhang, K. EMBO J. (1992) 11:4007-4015
#title Isolation of multiple mouse cDNAs with coding homology to *Saccharomyces cerevisiae* CDC25: identification of a region related to Bcr, Vav, Dbl and CDC24.

#cross-references M01D:93010996
#accession S28407
#status not compared with conceptual translation
##residues 1-1260 ##label CEN
##molecule_type mRNA

REFERENCE S22693
#authors Martegani, E.; Vannoni, M.; Zippel, R.; Coccetti, P.; Brambilla, R.; Ferrari, C.; Sturani, E.; Alberghina, L. EMBO J. (1992) 11:2151-2157
#title Cloning by functional complementation of a mouse cDNA encoding a homologue of CDC25, a *Saccharomyces cerevisiae* RAS activator.

#cross-references M01D:92289680
#accession S22693
#molecule_type mRNA
##residues 789-1260 ##label MAR
##cross-references EMBL:X59868; NID:g50357; PID:g50358
#note

REFERENCE A46199
#authors Wei, W.; Mosteller, R.D.; Sanyal, P.; Gonzales, E.; McKinney, D.; Dasgupta, C.; Li, P.; Liu, B.X.; Brock, D. Proc. Natl. Acad. Sci. U.S.A. (1992) 89:7100-7104
#title Identification of a mammalian gene structurally and functionally related to the CDC25 gene of *Saccharomyces cerevisiae*.

#cross-references M01D:92357779
#accession B46199
#status Preliminary
##residues 1029-1030, 'D', 1032-1224 ##label WEI
##experimental_source fetus
#note sequence extracted from NCBI backbone (NCBIN:111101, NCBI:P:111102)

CLASSIFICATION #superfamily CDC25-type guanine nucleotide exchange activator
#journal homology; CDC24 homology; pleckstrin repeat homology
FEATURE 242-428 #domain CDC24 homology #label CD24
1021-1257 #domain CDC25-type guanine nucleotide exchange activator homology #label SOS
SUMMARY #length 1260 #molecular-weight 143900 #checksum 9725

Query Match 77.3%; Score 51; DB 2; Length 1260;
Best Local Similarity 50.0%; Pred. No. 5,636+00;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 1058 EKYRTPY 1065
:::|:|:|

QY 1 QRYNRAPY 8

RESULT 7
ENTRY G02630 #type complete
TITLE FcalphaRb - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998

ACCESSIONS G02630
REFERENCE H01508
#authors van Dijk, T.B.; Morton, H.C.; Caldenhoven, E.; Bracke, M.;
#submission Raaijmakers, J.A.M.; Lammers, J.
#accession G02630
#status submitted to the EMBL Data Library, April 1996
#molecule_type mRNA
#residues 1-239 #label VAN
#cross-references EMBL:U56236; NID:g1326228; PID:g1326229

SUMMARY #length 239 #molecular_weight 26996 #checksum 5338

Query Match 75.8%; Score 50; DB 2; Length 239;
Best Local Similarity 71.4%; Pred. No. 8,806+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 196 WYNRSPY 202
:::|:|:|

QY 2 RYNRAPY 8

RESULT 8
ENTRY JH0332 #type complete
TITLE IGA (Fc) receptor, myeloid cell (CD89) precursor - human
ALTERNATE_NAMES myeloid glycoprotein CD89
ORGANISM #formal_name Homo sapiens #common_name man
DATE 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 10-Sep-1997

ACCESSIONS JH0332
REFERENCE JH0332
#authors Maliszewski, C.R.; March, C.J.; Schoenborn, M.A.; Gimpel, S.;
Shen, L.
#journal J. Exp. Med. (1990) 172:1665-1672
#title Expression cloning of a human Fc receptor for Iga.
#cross-references MUID:91079769
#accession JH0332

GENETICS
#residues 1-287 #label MAL
#cross-references GB:X54150; NID:g31329; PID:g31330
#experimental_source myeloid cell liver V937
#reference 137224
#authors de Wit, T.P.; Morton, H.C.; Capel, P.J.; van de Winkel, J.G.
#journal J. Immunol. (1995) 155:1203-1209
#title Structure of the gene for the human myeloid Iga Fc receptor (CD89).

ACCESSIONS #cross-references MUID:95363085
#accession I37224
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-287 #label RES
#cross-references EMBL:X87767; NID:g963041; PID:g1054737

GENETICS
#gene GDB:FCAR; CD89
#cross-references GDB:127543; OMIM:147045
#map_position 19q13.2-19q13.4
#introns 12/1; 24/1; 121/1; 217/1
KEYWORDS glycoprotein; immunoglobulin receptor; transmembrane protein

FEATURE
1-21 #domain signal sequence #status predicted #label SIG\

22-287 #product Iga receptor Fc alpha #status predicted #label MFC\

228-246 #domain transmembrane #status predicted #label TRA\

65,79,141,177,186 #binding-site carbohydrate (Asn) (covalent) #status predicted

SUMMARY #length 287 #molecular_weight 32265 #checksum 5963

Query Match 75.8%; Score 50; DB 2; Length 287;
Best Local Similarity 71.4%; Pred. No. 8,806+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 196 WYNRSPY 202
:::|:|:|

QY 2 RYNRAPY 8

RESULT 9
ENTRY I39956 #type complete
TITLE neutral proteinase (EC 3.4.24.-) - Bacillus amyloliquefaciens
ORGANISM #formal_name Bacillus amyloliquefaciens
DATE 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 18-Mar-1997

ACCESSIONS I39956
REFERENCE I39956
#authors Shimada, H.; Honjo, M.; Mita, I.; Nakayama, A.; Akaoka, A.;
Manabe, K.; Furutani, Y.
#journal J. Biotechnol. (1985) 2:75-85
#title The nucleotide sequence and some properties of the neutral
#accession I39956
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-521 #label RES
#cross-references GB:M36723; NID:g143352; PID:g143353

GENETICS
#start_codon GTG
CLASSIFICATION #superfamily thermolysin
KEYWORDS hydrolase; metalloproteinase
SUMMARY #length 521 #molecular_weight 56725 #checksum 6616

Query Match 75.8%; Score 50; DB 2; Length 521;
Best Local Similarity 62.5%; Pred. No. 8,806+00;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 308 OKYNRSY 315
:::|:|:|

QY 1 QRYNRAPY 8

RESULT 10
ENTRY S52662 #type complete
TITLE S-adenosylmethionine decarboxylase (SAMDC) - potato
ORGANISM #formal_name Solanum tuberosum #common_name potato
DATE 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 21-Jul-1995

ACCESSIONS S52662
REFERENCE S52662
#authors Arif, S.A.M.; Taylor, M.A.; George, L.A.; Butler, A.R.;
Burch, L.R.; Davies, H.V.; Stark, M.J.R.; Kumar, A.
#journal Plant Mol. Biol. (1994) 26:327-338
#title Characterisation of the S-adenosylmethionine decarboxylase (SAMDC) gene of potato.
#accession S52662
#status preliminary
#molecule_type DNA
#residues 1-360 #label ARI

SUMMARY #length 360 #molecular_weight 39724 #checksum 8045

Query Match 74.2%; Score 49; DB 2; Length 360;
Best Local Similarity 50.0%; Pred. No. 1,376+01;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 332 OKFRTPY 339
::: 1:1
QY 1 YNRNRPY 8

RESULT 11
ENTRY S28047 #type complete
TITLE TUB13 protein - potato
ORGANISM #formal_name Solanum tuberosum #common_name potato
DATE 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Sep-1997

ACCESSIONS
REFERENCE S28047
#authors Taylor, M.A.; Arif, S.A.M.; Kumar, A.; Davies, H.V.; Scobie, L.A.; Pearce, S.R.; Flavell, A.J.
#journal Plant Mol. Biol. (1992) 20:641-651
#title Expression and sequence analysis of cDNAs induced during the early stages of tuberisation in different organs of the potato plant (Solanum tuberosum L.).

GENETICS
#accession S28047
#molecule_type mRNA
#residues 1-360 #label TAY
#cross-references EMBL:Z11680; NID:921484; PID:921485

SUMMARY
#gene TUB13
#length 360 #molecular-weight 39726 #checksum 7941

Query Match 74.2%; Score 49; DB 2; Length 360;
Best Local Similarity 50.0%; Pred. No. 1.37e+01;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 332 OKFRTPY 339
::: 1:1
QY 1 YNRNRPY 8

RESULT 12
ENTRY S49009 #type complete
TITLE fork head protein 2 - African clawed frog
ORGANISM #formal_name Xenopus laevis #common_name African clawed frog
DATE 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 17-Mar-1999

ACCESSIONS
REFERENCE S49009
#authors Lef, J.; Clement, J.H.; Oschwald, R.; Koester, M.; Knoechel, W.
#journal Mech. Dev. (1994) 45:117-126
#title Spatial and temporal transcription patterns of the forkhead related XFD-2/XFD-2' genes in Xenopus laevis embryos.
#cross-references MUID:94257528
#accession S49009
#status preliminary
#molecule_type mRNA
#residues 1-367 #label LEF
#cross-references EMBL:X74316; NID:951161; PID:951162
CLASSIFICATION #superfamily_unassigned fork head proteins; fork head DNA-binding domain homology

FEATURE
128-219 #domain fork head DNA-binding domain homology #label FHD
SUMMARY
#length 367 #molecular-weight 40971 #checksum 6845

Query Match 72.7%; Score 48; DB 2; Length 367;
Best Local Similarity 83.3%; Pred. No. 2.11e+01;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 337 YNRSPY 342
::: 1:1
QY 3 YNRAPY 8

RESULT 13
ENTRY S49008 #type complete

TITLE fork head protein - African clawed frog
ORGANISM #formal_name Xenopus laevis #common_name African clawed frog
DATE 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 17-Mar-1999

ACCESSIONS
REFERENCE S49008; B56556
#authors Lef, J.; Clement, J.H.; Oschwald, R.; Koester, M.; Knoechel, W.
#journal Mech. Dev. (1994) 45:117-126
#title Spatial and temporal transcription patterns of the forkhead related XFD-2/XFD-2' genes in Xenopus laevis embryos.
#cross-references MUID:94257528
#accession S49008
#status preliminary
#molecule_type mRNA
#residues 1-370 #label LEF
#cross-references EMBL:X74315; NID:951159; PID:951160
REFERENCE A56556
#authors Knoechel, S.; Lef, J.; Clement, J.; Klocke, B.; Halle, S.; Koester, M.; Knoechel, W.
#journal Mech. Dev. (1992) 38:157-165
#title Activin A induced expression of a fork head related gene in posterior chordamesoderm (notochord) of Xenopus laevis embryos.
#cross-references MUID:93041288
#accession B56556
#status preliminary; not compared with conceptual translation
#molecule_type mRNA
#residues 118-228 #label KNO
#experimental_source gastrula
#note sequence extracted from NCBI backbone (NCBI:118178)
CLASSIFICATION #superfamily_unassigned fork head proteins; fork head DNA-binding domain homology

FEATURE
127-218 #domain fork head DNA-binding domain homology #label FHD
SUMMARY
#length 370 #molecular-weight 41388 #checksum 6542

Query Match 72.7%; Score 48; DB 2; Length 370;
Best Local Similarity 83.3%; Pred. No. 2.11e+01;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 335 YNRSPY 340
::: 1:1
QY 3 YNRAPY 8

RESULT 14
ENTRY S71800 #type fragment
TITLE transcription factor FAST-1 - African clawed frog (fragment)
ALTERNATE_NAMES forkhead activin signal transducer 1
ORGANISM #formal_name Xenopus laevis #common_name African clawed frog
DATE 06-Dec-1996 #sequence_revision 25-Apr-1997 #text_change 05-Dec-1997

ACCESSIONS
REFERENCE S71800
#authors Chen, X.; Rubock, M.J.; Whitman, M.
#journal Nature (1996) 383:691-696
#title A transcriptional partner for MAD proteins in TGF-beta signalling.
#cross-references MUID:97032727
#accession S71800
#status nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-534 #label CHE
#cross-references EMBL:U70980
CLASSIFICATION #superfamily_unassigned fork head proteins; fork head DNA-binding domain homology

KEYWORDS
FEATURE
117-209 DNA binding; transcription factor
SUMMARY
Query Match 72.7%; Score 48; DB 2; Length 534;

Best Local Similarity 50.0%; Pred. No. 2.11e+01;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 113 HRYXKPY 120
:|:|:|
QY 1 QRYNRAPY 8

RESULT 15 S74487 #type complete
ENTRY hypothetical protein sl11060 - *Synechocystis* sp. (strain PCC
TITLE 6803)

ORGANISM #formal_name *Synechocystis* sp.

DATE #variety PCC 6803
25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change
21-Aug-1998

ACCESSIONS S74487
REFERENCE S74322
#authors Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.;
Nakamura, Y.; Miyajima, N.; Hirose, M.; Sugita, M.;
Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.;
Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimo,
S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;
Yasuda, M.; Tabata, S.

#journal DNA Res. (1996) 3:109-136
#title Sequence analysis of the genome of the unicellular
cyanobacterium *Synechocystis* sp. PCC6803. II. Sequence
determination of the entire genome and assignment of
potential protein-coding regions.

#cross-references MUID:97061201

#accession S74487
#status nucleic acid sequence not shown; translation not shown

#molecule_type DNA

#residues 1-1032 #label KAN

#cross-references EMBL:D90899; GB:AB001339; NID:g1651650; PID:d1017372;
PID:g1651711

#note the nucleotide sequence was submitted to the EMBL Data
Library, June 1996

GENETICS

#start_codon GTG
SUMMARY #length 1032 #molecular_weight 117162 #checksum 9339

Query Match 72.7%; Score 48; DB 2; Length 1032;
Best Local Similarity 57.1%; Pred. No. 2.11e+01;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 708 RFDRDPY 714
:|:|:|
2 RYNRAPY 8

Search completed: Thu Sep 2 11:17:44 1999
Job time : 14 secs.

WISKEY (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSearch: protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Sep 2 11:16:54 1999; Maspar time 3.38 Seconds
56.561 Million cell updates/sec

Subular output not generated.

Title: >US-08-599-226-3

Description: (1-9) from US08599226.pep

Perfect Score: 66

Sequence: 1 ORYNRPYX 9

Scoring table:

PAM 150
Gap 15

Searched: 170751 seqs, 2126608 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

a-geneseq35
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39

Statistics: Mean 16.526; Variance 49.437; scale 0.334

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	66	100.0	9 27	W27585	Anti-TNF-alpha antilbo	1.22e+00
2	66	100.0	9 27	W27582	Anti-TNF-alpha antilbo	1.22e+00
3	66	100.0	107 27	W27568	Anti-TNF-alpha antilbo	1.22e+00
4	61	92.4	9 27	W27579	Anti-TNF-alpha antilbo	4.64e+00
5	61	92.4	9 27	W27575	Anti-TNF-alpha antilbo	4.64e+00
6	61	92.4	9 27	W27571	Anti-TNF-alpha antilbo	4.64e+00
7	58	87.9	9 27	W27584	Anti-TNF-alpha antilbo	1.02e+01
8	56	84.8	9 27	W27572	Anti-TNF-alpha antilbo	1.71e+01
9	54	81.8	263 4	R22957	Human proteasome comp	2.87e+01
10	54	81.8	269 1	R22666	Protein used to raise	2.87e+01
11	51	77.3	9 27	W27570	Anti-TNF-alpha antilbo	6.14e+01
12	51	77.3	9 27	W27576	Anti-TNF-alpha antilbo	6.14e+01
13	51	77.3	9 27	W27577	Anti-TNF-alpha antilbo	6.14e+01
14	51	77.3	9 27	W27574	Anti-TNF-alpha antilbo	6.14e+01
15	50	77.3	170 8	R42371	Prod. of ORFs of plas	6.14e+01
16	50	75.8	287 11	R59920	Human Fc-alpha-R.	7.89e+01

17	50	75.8	287 7	R34030	Fc-alpha-R.	7.89e+01
18	49	74.2	360 13	R75006	Tonato S-adenosyl-met	1.01e+02
19	47	71.2	9 27	W27582	Anti-TNF-alpha antilbo	1.66e+02
20	47	71.2	9 27	W27573	Anti-TNF-alpha antilbo	1.66e+02
21	47	71.2	419 28	W26744	S. carnosus nitrate r	1.66e+02
22	47	71.2	447 3	R12368	smg p23A GDP Dissocia	1.66e+02
23	46	69.7	9 27	W27578	Anti-TNF-alpha antilbo	2.12e+02
24	46	69.7	67 31	W28077	Amino acid sequence o	2.12e+02
25	45	68.2	246 13	R11131	Tyrosine kinase domai	2.70e+02
26	45	68.2	466 13	R11132	N-terminal truncated	2.70e+02
27	45	68.2	507 13	R11133	Cytoplasmic tyrosine k	2.70e+02
28	45	68.2	507 15	R84181	Megakaryocyte kinase	2.70e+02
29	45	68.2	520 1	P94617	Neutral protease enco	2.70e+02
30	45	68.2	521 1	P51009	Sequence of neutral p	2.70e+02
31	45	68.2	528 34	W64454	Human mark protein.	2.70e+02
32	45	68.2	557 2	R06370	Protein with lipase a	2.70e+02
33	45	68.2	563 2	R47577	Lipase of Geotrichum	2.70e+02
34	45	68.2	563 2	R10330	Gene product with lip	3.43e+02
35	44	66.7	9 27	W27580	Anti-TNF-alpha antilbo	3.43e+02
36	44	66.7	107 39	W86127	Murine 708 VI amino a	3.43e+02
37	44	66.7	107 39	W86129	Protein sequence of d	3.43e+02
38	44	66.7	287 19	W06702	Partial medium chain-	3.43e+02
39	44	66.7	400 10	R36281	Chitin-deacetylase.	3.43e+02
40	44	66.7	454 32	W58586	Human histamine H2 re	3.43e+02
41	44	66.7	517 20	W04558	Carnation ACC synthas	3.43e+02
42	44	66.7	529 36	W70501	Human G-protein adren	3.43e+02
43	44	66.7	2251 14	R71009	Human neuronal calciu	3.43e+02
44	44	66.7	2270 14	R71010	Human neuronal calciu	3.43e+02
45	44	66.7	2408 2	R24305	Translation of ORF 1	3.43e+02

ALIGNMENTS

RESULT 1
ID W27585 standard; peptide: 9 AA.
AC W27585;
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody light chain CDR3.
KW Human: tumour necrosis factor-alpha: TNF-alpha; antibody; CDR3;
KW light chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HIVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
PD 14-AUG-1997.
PE 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226.
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRW, Kaymakcalan Z, Labkovsky B,
PI Mankovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
PI WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 72, 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) light chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance) and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L229 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple

CC scleriosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 9 AA;

Query Match 100.0%; Score 66; DB 27; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.22e+00;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 grynrapy 8
| | | | | | |
QY 1 QRYNRAPY 8

RESULT 2
W27562 standard; peptide; 9 AA.
AC W27562:
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody light chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW light chain: complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
FT Key location/Qualifiers
FT Misc-difference 9 /label= Thr, Ala
PN W09729131-A1.
PD 14-AUG-1997.
PF 10-FEB-1997: US-02219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226. 1 m s T
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
WPI: 97-415302/38.
CC High affinity antibodies against human TNF alpha - useful to inhibit
CC TNF alpha activity, e.g. to treat autoimmune diseases and cancer
CC Claim 9; Page 64; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) light chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 9 AA;

Query Match 100.0%; Score 66; DB 27; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.22e+00;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 grynrapy 8
| | | | | | |
QY 1 QRYNRAPY 8

RESULT 3
W27568 standard; Protein; 107 AA.
AC W27568:
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody light chain variable region.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody;
KW light chain: variable region; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
FT Key location/Qualifiers
FT Misc-difference 9 /label= Thr, Ala
PN W09729131-A1.
PD 14-AUG-1997.
PF 10-FEB-1997: US-02219.
PR 25-NOV-1996: US-031476. 1 m s T
PR 09-FEB-1996: US-599226.
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
WPI: 97-415302/38.
DR N-PSDB; T88403.
CC High affinity antibodies against human TNF alpha - useful to inhibit
CC TNF alpha activity, e.g. to treat autoimmune diseases and cancer
CC Claim 15; Page 75; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) light chain variable region.
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 107 AA;

Query Match 100.0%; Score 66; DB 27; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.22e+00;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 89 grynrapy 96
| | | | | | |
QY 1 QRYNRAPY 8

RESULT 4
W27579 standard; peptide; 9 AA.
AC W27579:

DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody light chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW light chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997: U02219.
 PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B, Mankovich JA, McGuinness BT, Roberts AJ, Sakorafas P, Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ; WPI: 97-415302/38.
 DR High affinity antibodies against human TNF alpha - useful to inhibit TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20: Page 70: 102pp: English.
 CC The present sequence is a novel anti-human tumour necrosis factor-alpha (TNF-alpha) antibody (Ab) light chain complementarity determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or less and has a Koff rate constant of 1x10 power -3 s power -1 or less (both determined by surface plasmon resonance), and neutralises human TNF-alpha cytotoxicity in a standard in vitro 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which inhibits TNF-alpha activity, can be used to treat sepsis, autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid spondylitis, osteoarthritis, gouty arthritis, allergy, multiple sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic syndrome, infectious diseases, malignancy, pulmonary, intestinal, cardiac or inflammatory bone disorders, bone resorption disease, alcoholic, viral or fulminant hepatitis, coagulation disturbances, burns, reperfusion injury, keloid formation, scar tissue formation, pyrexia, periodontal disease, obesity and radiation toxicity. The Ab also inhibits TNF-alpha induced expression of endothelial cell leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein endothelial cells (HUVEC).
 CC Sequence 9 AA:
 SQ

Query Match 92.4%; Score 61; DB 27; Length 9;
 Best Local Similarity 87.5%; Pred. NO. 4.64e+00;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 qkynraby 8
 I:|||||
 QY 1 QRYNRAPY 8

RESULT 5
 ID W27575 standard; peptide: 9 AA.
 DT 19-MAR-1998 (first entry)
 AC W27575;
 DE Anti-TNF-alpha antibody light chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW light chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.

PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997: U02219.
 PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B, Mankovich JA, McGuinness BT, Roberts AJ, Sakorafas P, Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ; WPI: 97-415302/38.
 DR High affinity antibodies against human TNF alpha - useful to inhibit TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20: Page 69: 102pp: English.
 CC The present sequence is a novel anti-human tumour necrosis factor-alpha (TNF-alpha) antibody (Ab) light chain complementarity determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or less and has a Koff rate constant of 1x10 power -3 s power -1 or less (both determined by surface plasmon resonance), and neutralises human TNF-alpha cytotoxicity in a standard in vitro 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which inhibits TNF-alpha activity, can be used to treat sepsis, autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid spondylitis, osteoarthritis, gouty arthritis, allergy, multiple sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic syndrome, infectious diseases, malignancy, pulmonary, intestinal, cardiac or inflammatory bone disorders, bone resorption disease, alcoholic, viral or fulminant hepatitis, coagulation disturbances, burns, reperfusion injury, keloid formation, scar tissue formation, pyrexia, periodontal disease, obesity and radiation toxicity. The Ab also inhibits TNF-alpha induced expression of endothelial cell leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein endothelial cells (HUVEC).
 CC Sequence 9 AA:
 SQ

Query Match 92.4%; Score 61; DB 27; Length 9;
 Best Local Similarity 87.5%; Pred. NO. 4.64e+00;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 qkynraby 8
 I:|||||
 QY 1 QRYNRAPY 8

RESULT 6
 ID W27571 standard; peptide: 9 AA.
 DT 19-MAR-1998 (first entry)
 AC W27571;
 DE Anti-TNF-alpha antibody light chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW light chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997: U02219.
 PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B, Mankovich JA, McGuinness BT, Roberts AJ, Sakorafas P, Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ; WPI: 97-415302/38.
 DR High affinity antibodies against human TNF alpha - useful to inhibit TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20: Page 68: 102pp: English.

CC The present sequence is a novel anti-human tumour necrosis
CC factor- α (TNF- α) antibody (Ab) light chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF- α with a K_d of 1×10^{-8} M or
CC less and has a Koff rate constant of 1×10^3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF- α cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1×10^3 pM or less. The Ab, which
CC inhibits TNF- α activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF- α induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
Sequence 9 AA:

Query Match 92.4%; Score 61; DB 27; Length 9;
Best Local Similarity 87.5%; Pred. No. 4.64e+00;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 qkynrpy 8
1:|||||
QY 1 QRYNRPY 8

RESULT 7
ID W2584 standard; peptide; 9 AA.

AC W2584;
DE Anti-TNF- α antibody light chain CDR3.
KW Human: tumour necrosis factor- α ; TNF- α ; antibody; CDR3;
KW light chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W0929131-A1.
PD 14-AUG-1997;
PF 10-FEB-1997; U02219;
PR 25-NOV-1996; US-031476.
PR 09-FEB-1996; US-599226.
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,
PI Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF α - useful to inhibit
PT TNF α activity, e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 72; 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor- α (TNF- α) antibody (Ab) light chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF- α with a K_d of 1×10^{-8} M or
CC less and has a Koff rate constant of 1×10^3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF- α cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1×10^3 pM or less. The Ab, which
CC inhibits TNF- α activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,

CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF- α induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
Sequence 9 AA:

Query Match 87.9%; Score 58; DB 27; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.03e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 qkynrpy 8
1:|||||
QY 1 QRYNRPY 8

RESULT 8
ID W2572 standard; peptide; 9 AA.

AC W2572;
DE Anti-TNF- α antibody light chain CDR3.
KW Human: tumour necrosis factor- α ; TNF- α ; antibody; CDR3;
KW light chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W0929131-A1.
PD 14-AUG-1997;
PF 10-FEB-1997; U02219;
PR 25-NOV-1996; US-031476.
PR 09-FEB-1996; US-599226.
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,
PI Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF α - useful to inhibit
PT TNF α activity, e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 68; 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor- α (TNF- α) antibody (Ab) light chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF- α with a K_d of 1×10^{-8} M or
CC less and has a Koff rate constant of 1×10^3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF- α cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1×10^3 pM or less. The Ab, which
CC inhibits TNF- α activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF- α induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
Sequence 9 AA:

Query Match 84.8%; Score 56; DB 27; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.71e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1 qkygrrpy 8
|:| | | |
QY 1 QRYNRPY 8

RESULT 9
ID R22957 standard; Protein: 263 AA.

AC R22957;
DT 15-OCT-1992 (first entry)
DE Human proteasome component HC2.
KW HC2; HC5; HC8; HC9; Alzheimer's disease; cancer.
OS Homo sapiens.

PN J04077498-A.
PD 11-MAR-1992.
PF 20-JUL-1990; 193313.
PR 20-JUL-1990; JP-193313.
PS (SARA) OTSUKA PHARM KK.
DR WPI: 92-136767/17.
DR N-NSDB: Q23894.

Human proteasome - has specified aminoacid sequence and gene base sequence, used for e.g. investigation of Alzheimer's disease
Claim 1: Page 1; 24pp; Japanese.
Human proteasome HC2 and its gene can be used in the investigation, diagnosis, and treatment of diseases associated with proteasome abnormality such as cancers and Alzheimer's disease.
Human proteasome is purified from the soluble liver fractions of human liver cell HepG2 cell. A probe was prepared, and a cDNA library constructed using plasmid Bluescript KS+. A clone was isolated and its DNA sequenced to give Q23894.
See also Q23894-7, R22957-60.
Sequence 263 AA;

Query Match 81.8%; Score 54; DB 4; Length 263;
Best Local Similarity 75.0%; Pred. No. 2.87e+01;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 121 qkygrrpy 128
|:| | | |
QY 1 QRYNRPY 8

RESULT 10
ID R22666 standard; Protein: 269 AA.
AC R22666;
DT 06-NOV-1992 (first entry)
DE Protein used to raise anti-p33k antibodies.
KW HIV; human immunodeficiency virus; promosomal; cell surface; p27K.
OS Homo sapiens.
PN MO9207269-A.
PD 30-APR-1992.
PF 10-OCT-1991; E01945.
PR 11-OCT-1990; EP-402838.
PS (PROS-) PRO-SOMA SARL.
PI Bey F, Bureau JP, Scherrer K;
DR WPI: 92-167288/20.
DR N-PSDB: Q24128.
PT Method for diagnosing HIV infection - comprises using immunochemical reagent contg. monoclonal antibodies against PT promosomal cell surface proteins.
PS Claim 13; Fig 9; 34pp; English.
CC The 269 amino acid sequence is used to raise antibodies against the CC this protein can detect promosomal surface antigens. HIV infected cells CC express these proteins on their surface, thus the antibodies may be used CC to detect HIV infection and to determine whether T4 cells have been CC infected even before they are killed by HIV.
CC See also R22665.
SO Sequence 269 AA;

Query Match 81.8%; Score 54; DB 1; Length 269;
Best Local Similarity 75.0%; Pred. No. 2.87e+01;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 127 qkygrrpy 134
|:| | | |
QY 1 QRYNRPY 8

RESULT 11
ID W27570 standard; peptide: 9 AA.

AC W27570;
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody light chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW light chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.

OS Homo sapiens.
PN W09729131-A1.
PD 14-AUG-1997.
PF 10-FEB-1997; U02219.
PR 25-NOV-1996; US-031476.
PS 09-FEB-1996; US-599226.
PA (RADT) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Menkovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Willton AJ;
DR WPI: 97-415302/38.

High affinity antibodies against human TNF alpha - useful to inhibit TNF alpha activity, e.g. to treat autoimmune diseases and cancer
Claim 20; Page 67; 102pp; English.
The present sequence is a novel anti-human tumour necrosis factor-alpha (TNF-alpha) antibody (Ab) light chain complementarity determining region 3 (CDR3).
The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or less and has a Koff rate constant of 1x10 power -3 s power -1 or less (both determined by surface plasmon resonance), and neutralises human TNF-alpha cytotoxicity in a standard in vitro 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which inhibits TNF-alpha activity, can be used to treat sepsis, autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid spondylitis, osteoarthritis, gouty arthritis, allergy, multiple sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic syndrome, infectious diseases, malignancy, pulmonary, intestinal, cardiac or inflammatory bone disorders, bone resorption disease, CC alcoholic, viral or fulminant hepatitis, coagulation disturbances, CC burns, reperfusion injury, keloid formation, scar tissue formation, CC pyrexia, periodontal disease, obesity and radiation toxicity. The CC Ab also inhibits TNF-alpha induced expression of endothelial cell CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein CC endothelial cells (HUVEC).
SO Sequence 9 AA;

Query Match 77.3%; Score 51; DB 27; Length 9;
Best Local Similarity 75.0%; Pred. No. 6.14e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 1 qkygrrpy 8
|:| | | |
QY 1 QRYNRPY 8

RESULT 12
ID W27576 standard; peptide: 9 AA.

AC W27576;
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody light chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW light chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;

KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN WO9729131-A1.
PD 14-AUG-1997.
PF 10-FEB-1997: US-031476.
PR 25-NOV-1996: US-599226.
PI (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Markovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
TNF alpha activity, e.g. to treat autoimmune diseases and cancer
Claim 20: Page 69; 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
factor-alpha (TNF-alpha) antibody (Ab) light chain complementarity
determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
less and has a Koff rate constant of 1x10 power -3 s power -1 or
less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SO Sequence 9 AA:

Query Match 77.3%; Score 51; DB 27; Length 9;
Best Local Similarity 75.0%; Pred. No. 6.14e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 qkynsapy 8
1 ORYNKAPR 8

RESULT 13
ID W27577 standard; peptide: 9 AA.

AC W27577;
DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody light chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW light chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN WO9729131-A1.
PD 14-AUG-1997.
PF 10-FEB-1997: US-031476.
PR 25-NOV-1996: US-599226.
PI (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Markovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
TNF alpha activity, e.g. to treat autoimmune diseases and cancer
Claim 20: Page 69; 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
factor-alpha (TNF-alpha) antibody (Ab) light chain complementarity
determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
less and has a Koff rate constant of 1x10 power -3 s power -1 or
less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SO Sequence 9 AA:

Query Match 77.3%; Score 51; DB 27; Length 9;
Best Local Similarity 75.0%; Pred. No. 6.14e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 qkynsapy 8
1 ORYNKAPR 8

PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Markovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
TNF alpha activity, e.g. to treat autoimmune diseases and cancer
Claim 20: Page 70; 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
factor-alpha (TNF-alpha) antibody (Ab) light chain complementarity
determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
less and has a Koff rate constant of 1x10 power -3 s power -1 or
less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SO Sequence 9 AA:

Query Match 77.3%; Score 51; DB 27; Length 9;
Best Local Similarity 75.0%; Pred. No. 6.14e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 qkynsapy 8
1 ORYNKAPR 8

RESULT 14
ID W27574 standard; peptide: 9 AA.

AC W27574;
DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody light chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW light chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN WO9729131-A1.
PD 14-AUG-1997.
PF 10-FEB-1997: US-031476.
PR 25-NOV-1996: US-599226.
PI (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Markovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
TNF alpha activity, e.g. to treat autoimmune diseases and cancer
Claim 20: Page 69; 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
factor-alpha (TNF-alpha) antibody (Ab) light chain complementarity
determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
less and has a Koff rate constant of 1x10 power -3 s power -1 or
less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SO Sequence 9 AA:

Query Match 77.3%; Score 51; DB 27; Length 9;
Best Local Similarity 75.0%; Pred. No. 6.14e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 qkynsapy 8
1 ORYNKAPR 8

CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 9 AA;

Query Match 77.3%; Score 51; DB 27; Length 9;
 Best Local Similarity 75.0%; Pred. No. 6.14e+01;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 1 qkynsapy 8
 1:11 111
 QY 1 ORYNRAPY 8

QY 1 ORYNRAPY 8

Search completed: Thu Sep 2 11:17:12 1999
 Job time : 18 secs.

RESULT 15
 ID R42371 standard; Protein: 170 AA.
 AC R42371;
 DT 19-APR-1994 (first entry)
 DE Prod. of ORFs of plasmid PRAP501.
 KW Haemophilus somnus; immunogenic; haemolysin; Lppb; Lppc;
 KW thromboembolic meningoencephalitis; septicemia; arthritis;
 KW pneumonia; haemin-binding protein.
 OS Haemophilus somnus.
 PN WO9321323-A.
 PD 28-OCT-1993.
 PE 05-APR-1993; CA0135.
 PR 09-APR-1992; US-865050.
 PR 04-JUN-1992; US-893424.
 PR 04-JUN-1992; US-893426.
 PR 29-MAR-1993; US-038287.
 PR 29-MAR-1993; US-038288.
 PR 29-MAR-1993; US-038719.
 PA (USRA-) UNIV SASKATCHEWAN.
 PI Harland RJ, Pfeiffer CG, Pontarollo RA, Potter AA;
 PI Rioux C, Theisen M;
 PI WPI: 93-351733/44.
 PI N-PSDB: 051080.
 PI Haemophilus somnus immunogenic proteins used in vaccines -
 PI selected from haemin-binding protein, haemolysin, Lppb and Lppc,
 PI and corresp. DNA
 PS Disclosure: Fig 2; 119pp; English.
 CC A genomic cosmid library of Haemophilus somnus HS25 DNA was screened
 CC for clones capable of binding bovine haemin and having haemolytic
 CC activity. Positive clones were subcloned various times, resulting
 CC in plasmid PRAP501, which binds haemin but is not haemolytic. The
 CC clone was sequenced and was found to contain several open reading
 CC frames, potentially encoding 8 proteins. The haemin binding protein
 CC (encoded by the hmb gene) was encoded by ORF1. The protein can be
 CC used in vaccines for preventing or treating H. somnus infections,
 CC which cause thromboembolic meningo-encephalitis, septicemia,
 CC arthritis and pneumonia in vertebrates. The protein shown is
 CC potentially encoded by ORF5, and has an unknown function.
 CC See also R42370-86.
 SQ Sequence 170 AA;

Query Match 77.3%; Score 51; DB 8; Length 170;
 Best Local Similarity 75.0%; Pred. No. 6.14e+01;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 107 qkynsapy 114
 11:11 1

THIS PAGE BLANK (USPTO)

 WISE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPearch_pp protein - protein database search, using Smith-Waterman algorithm

on: Thu Sep 2 11:20:18 1999; MasPar time 3.47 Seconds
 73.528 Million cell updates/sec

Title: >US-08-599-226-4
 Description: (1-12) from US08599226.pep
 Perfect Score: 66

Sequence: 1 VSYLSTASLDX 12

Scoring table: PAM 150
 Gap 15

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

a-geneseg35
 1:part1 2:part2 3:part3 4:part4 5:parts 6:part6 7:part7
 8:part8 9:part10 11:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 16.837; Variance 49.496; scale 0.340

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	66	100.0	12 27	W27594	Anti-TNF-alpha antibo	1.22e+00
2	66	100.0	12 27	W27594	Anti-TNF-alpha antibo	1.22e+00
3	66	100.0	12 27	W27594	Anti-TNF-alpha antibo	1.22e+00
4	58	87.9	12 27	W27593	Anti-TNF-alpha antibo	1.04e+01
5	57	86.4	12 27	W27586	Anti-TNF-alpha antibo	1.35e+01
6	57	86.4	12 27	W27587	Anti-TNF-alpha antibo	1.35e+01
7	57	86.4	12 27	W27588	Anti-TNF-alpha antibo	1.35e+01
8	57	86.4	12 27	W27589	Anti-TNF-alpha antibo	1.35e+01
9	52	78.8	12 27	W28236	Amino acid sequence o	4.92e+01
10	51	77.3	12 27	W27592	Anti-TNF-alpha antibo	6.34e+01
11	49	74.2	12 27	W27591	Anti-TNF-alpha antibo	1.05e+02
12	48	72.7	12 27	W27590	Anti-TNF-alpha antibo	1.35e+02
13	48	72.7	12 27	W23067	Canine IgE heavy chai	1.35e+02
14	48	72.7	12 27	W40054	P300/CBP-associated t	1.35e+02
15	48	72.7	12 27	W40052	Human P300/CBP-associ	1.35e+02
16	48	72.7	12 27	W81867	Human tumour suppress	1.35e+02

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
17	47	71.2	9 12	R61854	MBP peptide 14, poten	1.73e+02
18	47	71.2	10 12	R61948	MBP peptide 14, poten	1.73e+02
19	47	71.2	11 12	R61957	MBP peptide 14, poten	1.73e+02
20	47	71.2	13 16	R85136	Human MBP residues 13	1.73e+02
21	47	71.2	19 19	R95335	MBP-1.1 (11-29)	1.73e+02
22	47	71.2	19 27	W37551	Human myelin basic pr	1.73e+02
23	47	71.2	20 39	W73609	Human myelin basic pr	1.73e+02
24	47	71.2	20 39	R95334	MBP-1 (11-30)	1.73e+02
25	47	71.2	20 27	W37550	Human myelin basic pr	1.73e+02
26	47	71.2	21 35	W78823	Human myelin basic pr	1.73e+02
27	47	71.2	21 19	R95336	MBP-1.2 (11-31)	1.73e+02
28	47	71.2	127 9	R48596	Rat myelin basic prot	1.73e+02
29	47	71.2	142 35	W38731	Streptococcus pneumon	1.73e+02
30	47	71.2	167 9	R48595	Myelin basic protein	1.73e+02
31	47	71.2	168 9	R48594	Rabbit myelin basic p	1.73e+02
32	47	71.2	168 1	R04717	Empirically determine	1.73e+02
33	47	71.2	170 6	R30736	Human MBP	1.73e+02
34	47	71.2	170 19	R95406	Myelin Oligodendrocyt	1.73e+02
35	47	71.2	170 7	R35440	Human basic myelin pr	1.73e+02
36	47	71.2	170 9	R48592	Human myelin basic pr	1.73e+02
37	47	71.2	171 18	R97627	Human myelin basic pr	1.73e+02
38	47	71.2	171 18	R99580	Human myelin basic pr	1.73e+02
39	47	71.2	197 19	W00399	Human myelin basic pr	1.73e+02
40	47	71.2	203 19	W06108	Foetal myelin basic p	1.73e+02
41	47	71.2	203 19	W06107	Foetal myelin basic p	1.73e+02
42	47	71.2	373 19	W06103	MP4 chimera (MBP21.5-	1.73e+02
43	47	71.2	375 19	W06104	PM4 chimera (delta PL	1.73e+02
44	47	71.2	385 19	W06102	MP3 chimera (MBP21.5-	1.73e+02
45	47	71.2	492 19	W06105	MOGp4 chimera (MBP21	1.73e+02

ALIGNMENTS

RESULT 1
 ID W27594 standard; peptide: 12 AA.
 AC W27594;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HIVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN W09729131-A1.
 PD 14-AUG-1997.
 PE 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankevich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WPI: 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Disclosure: Page 75; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -3 s power -1 or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple

CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

Query Match 100.0%; Score 66; DB 27; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.22e+00;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 vsy1stassid 11
| | | | | | | | | |
QY 1 VSY1STASSLD 11

RESULT 2
W27563 standard; peptide; 12 AA.

AC W27563;
DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
FT Key Location/Qualifiers
FT Misc-difference 12 /label: Tyr, Asn
PD W09729131-A1.
PE 14-AUG-1997.
PF 10-FEB-1997; U02219.
PR 25-NOV-1996; US-031476.
PR 09-FEB-1996; US-599226.
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
PI Saifeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
PI WPI: 97-415302/38
CC High affinity antibodies against human TNF alpha - useful to inhibit
CC TNF alpha activity, e.g. to treat autoimmune diseases and cancer
CC Claim 9; Page 65; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Kof rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

Query Match 100.0%; Score 66; DB 27; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.22e+00;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 vsy1stassid 11
| | | | | | | | | |
QY 1 VSY1STASSLD 11

RESULT 3

W27569 standard; protein; 121 AA.

AC W27569;
DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain variable region.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody;
KW heavy chain; variable region; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
FT Key Location/Qualifiers
FT Misc-difference 12 /label: Tyr, Asn
PD W09729131-A1.
PE 14-AUG-1997.
PF 10-FEB-1997; U02219.
PR 25-NOV-1996; US-031476.
PR 09-FEB-1996; US-599226.
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
PI Saifeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
PI WPI: 97-415302/38.
DR N-FSDS; T88404.

FT High affinity antibodies against human TNF alpha - useful to inhibit
FT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
FT Claim 16; Page 76; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain variable region.
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Kof rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 121 AA;

Query Match 100.0%; Score 66; DB 27; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.22e+00;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 99 vsy1stassid 109
| | | | | | | | | |
QY 1 VSY1STASSLD 11

RESULT 4
W27593 standard; peptide; 12 AA.
AC W27593;

DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226. *ms*
 PA (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR MPI: 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 73; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical cell
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA;
 SQ

Query Match 87.9%; Score 58; DB 27; Length 12;
 Best Local Similarity 90.0%; Pred. No. 1.04e+01;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 2 systassle 11
 |||||:
 QY 2 SYLSTASSLD 11

RESULT 5
 ID W27586 standard: peptide: 12 AA.
 AC W27586;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.

PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226. *ms*
 PA (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR MPI: 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 72; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA;
 SQ

Query Match 86.4%; Score 57; DB 27; Length 12;
 Best Local Similarity 90.0%; Pred. No. 1.35e+01;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 2 systassle 11
 |||||:
 QY 2 SYLSTASSLD 11

RESULT 6
 ID W27587 standard: peptide: 12 AA.
 AC W27587;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226. *ms*
 PA (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR MPI: 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 73; 102pp; English.

CC The present sequence is a novel anti-human tumour necrosis
CC factor- α (TNF- α) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF- α with a K_d of 1×10^9 power -8 M or
CC less and has a Koff rate constant of 1×10^9 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF- α cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1×10^9 power -7 M or less. The Ab, which
CC inhibits TNF- α activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF- α induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
Sequence 12 AA:

Query Match 86.4%: Score 57: DB 27: Length 12:
Best Local Similarity 90.0%: Pred. No. 1.35e+01:
Matches 9: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Db 2 sylstssald 11
|||||
OY 2 SYLSTASSLD 11

RESULT 7
ID W27589 standard; peptide; 12 AA.

AC W27589:
DT 19-MAR-1998 (first entry)
DE Anti-TNF- α antibody heavy chain CDR3.
KW Human; tumour necrosis factor- α ; TNF- α ; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
RN M09729131-A1.
PR 14-AUG-1997.
PR 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226.
PI (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankovich JA, McGulness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF α - useful to inhibit
PT TNF α activity, e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 73; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor- α (TNF- α) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF- α with a K_d of 1×10^9 power -8 M or
CC less and has a Koff rate constant of 1×10^9 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF- α cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1×10^9 power -7 M or less. The Ab, which
CC inhibits TNF- α activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,

CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF- α induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
Sequence 12 AA:

Query Match 86.4%: Score 57: DB 27: Length 12:
Best Local Similarity 90.0%: Pred. No. 1.35e+01:
Matches 9: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Db 2 sylstssald 11
|||||
OY 2 SYLSTASSLD 11

RESULT 8
ID W27588 standard; peptide; 12 AA.

AC W27588:
DT 19-MAR-1998 (first entry)
DE Anti-TNF- α antibody heavy chain CDR3.
KW Human; tumour necrosis factor- α ; TNF- α ; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
RN M09729131-A1.
PR 14-AUG-1997.
PR 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226.
PI (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankovich JA, McGulness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF α - useful to inhibit
PT TNF α activity, e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 73; 102pp; English.

CC The present sequence is a novel anti-human tumour necrosis
CC factor- α (TNF- α) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF- α with a K_d of 1×10^9 power -8 M or
CC less and has a Koff rate constant of 1×10^9 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF- α cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1×10^9 power -7 M or less. The Ab, which
CC inhibits TNF- α activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF- α induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
Sequence 12 AA:

Query Match 86.4%: Score 57: DB 27: Length 12:
Best Local Similarity 90.0%: Pred. No. 1.35e+01:
Matches 9: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Db 2 systssslid 11
|:|:|:|:|:|
QY 2 SYLSTASSLD 11

RESULT 9
ID W28236 standard; Protein: 441 AA.
AC W28236;

DE 07-SEP-1998 (first entry)
KW Amino acid sequence of a mercuric reductase.
KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW Staphylococcal gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
KW toxic shock syndrome; mercuric reductase.
OS Staphylococcus aureus.

FT key location/Qualifiers
FT Misc-difference 16
PN WO9730070-A1.
21-AUG-1997.
19-FEB-1997: 002318.
20-FEB-1996: US-011888.

PA (SMIK) SMITHKLINE BEECHAM CORP.
PI Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO,
PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;
DR WPI: 97-424569/39.
N-PSDB: T84151.

PT Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used
PT to isolate antimicrobial compounds, and in vaccines against S.

PS aureus infection

CC Claim 6: Pages 547-548, 989pp; English.

CC The present sequence represents a Staphylococcus aureus protein, that,
CC based on homology with another Staphylococcus aureus protein, is
CC believed to be a mercuric reductase. The DNA sequence was isolated from
CC a library of clones of S. aureus WCUH 29 in Escherichia coli. The DNA

CC sequence can be used in the construction of ribozymes and antisense
CC sequences to control the expression of Staphylococcal genes. The DNA
CC sequence is also useful as a source of regulatory elements for the
CC control of bacterial gene expression. The present protein may be used
CC to produce vaccines to enable a host to produce specific antibodies
CC with antibacterial action. These vaccines and antibodies would protect
CC a host against invasion by S. aureus, and conditions relating to
CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled
CC skin syndrome, and toxic shock syndrome.

QY Sequence 441 AA;

Query Match 78.8%; Score 52; DB 32; Length 441;

Best Local Similarity 45.5%; Pred. No. 4.92e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 203 vdytstssale 213
|:|:|:|:|:|
QY 1 VSYLSTASSLD 11

RESULT 10
ID W27592 standard; peptide: 12 AA.
AC W27592;

DE 19-MAR-1998 (first entry)

KW Anti-TNF-alpha antibody heavy chain CDR3.

KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;

KW heavy chain; complementarity determining region 3; inhibition;

KW treatment; sepsis; disease; autoimmune disease; infectious disease;

KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;

KW cardiac disorder; inflammatory bone disorder; reperfusion injury;

KW bone resorption disease; coagulation disturbance; burn; ELAM-1;

KW keloid formation; scar tissue formation; pyrexia; HUVEC;

KW periodontal disease; obesity; radiation toxicity;

KW endothelial cell leukocyte adhesion molecule-1;

KW human umbilical vein endothelial cell.

OS Homo sapiens.

PN WO9729131-A1.
14-AUG-1997.

PE 10-FEB-1997: 002219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226.
PA (BADI) BASF AG.

PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit

PS Claim 20, Page 74; 102pp; English.

CC The present sequence is a novel anti-human tumour necrosis

CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity

CC determining region 3 (CDR3).

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or

CC less and has a Koff rate constant of 1x10 power -3 s power -1 or

CC less (both determined by surface plasmon resonance), and

CC neutralises human TNF-alpha cytotoxicity in a standard in vitro

CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which

CC inhibits TNF-alpha activity, can be used to treat sepsis,

CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid

CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple

CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic

CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,

CC cardiac or inflammatory bone disorders, bone resorption disease,

CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,

CC burns, reperfusion injury, keloid formation, scar tissue formation,

CC pyrexia, periodontal disease, obesity and radiation toxicity. The

CC Ab also inhibits TNF-alpha induced expression of endothelial cell

CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein

CC endothelial cells (HUVEC).

QY Sequence 12 AA;

Query Match 77.3%; Score 51; DB 27; Length 12;
Best Local Similarity 70.0%; Pred. No. 6.34e+01;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 2 sflstssale 11
|:|:|:|:|:|
QY 2 SYLSTASSLD 11

RESULT 11
ID W27591 standard; peptide: 12 AA.
AC W27591;

DE 19-MAR-1998 (first entry)

KW Anti-TNF-alpha antibody heavy chain CDR3.

KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;

KW heavy chain; complementarity determining region 3; inhibition;

KW treatment; sepsis; disease; autoimmune disease; infectious disease;

KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;

KW cardiac disorder; inflammatory bone disorder; reperfusion injury;

KW bone resorption disease; coagulation disturbance; burn; ELAM-1;

KW keloid formation; scar tissue formation; pyrexia; HUVEC;

KW periodontal disease; obesity; radiation toxicity;

KW endothelial cell leukocyte adhesion molecule-1;

KW human umbilical vein endothelial cell.

OS Homo sapiens.

PN WO9729131-A1.
14-AUG-1997.

DE 10-FEB-1997: 002219.

PR 25-NOV-1996: US-031476.

PR 09-FEB-1996: US-599226.

PA (BADI) BASF AG.

PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit

PS Claim 20, Page 74; 102pp; English.

CC The present sequence is a novel anti-human tumour necrosis

CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity

CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis.
 CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:

Query Match 77.3%; Score 51; DB 27; Length 12;
 Best Local Similarity 88.9%; Pred. No. 6.34e+01;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 2 sylstsssl 10
 QY 2 sylstsssl 10

RESULT 12
 ID W27590 standard; peptide; 12 AA.

AC W27590;
 DT 19-MAR-1998 (first entry)
 KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN W09729131-A1.
 PD 14-AUG-1997.
 RF 10-FEB-1997; U02219.
 RF 25-NOV-1996; US-031476.
 RF 09-FEB-1996; US-599226.
 RA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Markovitch JA, McGuinness BF, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR WPI: 97-415302/38.
 FT High affinity antibodies against human TNF alpha - useful to inhibit
 FT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis.
 CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,

CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:

Query Match 74.2%; Score 49; DB 27; Length 12;
 Best Local Similarity 80.0%; Pred. No. 1.05e+02;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 2 sylstsssl 11
 QY 2 sylstsssl 11

RESULT 13
 ID W23067 standard; Protein; 417 AA.

AC W23067;
 DT 19-FEB-1998 (first entry)
 KW Canine IgE heavy chain constant region (exon 1-4 product).
 KW IgE; immunoglobulin; antibody; heavy chain constant region;
 KW allergy; hypersensitivity; therapy; dog; antisense;
 KW immunomodulation.
 OS Canis familiaris.
 FH Key
 FT MISC_difference 55 Location/Qualifiers
 FT FT /note= "encoded by ACC"
 FT MISC_difference 56 /note= "encoded by TAC"
 FT FT /note= "encoded by TAC"
 FT MISC_difference 67 /note= "encoded by GCC"
 FT FT /note= "encoded by NNT"
 FT MISC_difference 174 /note= "encoded by GGN"
 FT MISC_difference 175 /note= "encoded by NNG"
 FT MISC_difference 176 /note= "encoded by TGN"
 FT MISC_difference 203 /note= "encoded by TCC"
 FT MISC_difference 204 /note= "encoded by GAC"
 FT FT /note= "encoded by GAC"
 PN W09730156-A2.
 PD 21-AUG-1997.
 PE 14-FEB-1997; U02322.
 PR 14-FEB-1996; US-601197.
 RA (IDEX-) IDEX LAB INC.
 PI Harris RA, Mermer B, Sleifing AE;
 DR WPI: 97-425031/39.
 DR N-PSDB: T79278.
 PT Isolated canine IgE heavy chain constant region DNA - useful to
 PT develop products for treatment of canine allergies and for
 PT immunomodulation in dogs
 PS Disclosure; Page 35-39; 59pp; English.
 CC This polypeptide is encoded by exons 1-4 (see T79278) of canine
 CC IgE heavy chain constant region (epsilon) genomic DNA. Another
 CC polypeptide, comprising the exon 5 and 6 product, is given in
 CC W23068. Recombinant peptides encoded by exons 1-6 can be
 CC produced in eukaryotic or prokaryotic cells. Such peptides,
 CC and antibodies raised against them, are used in methods to treat
 CC the manifestation of allergy in dogs, e.g. to treatment Type I
 CC immediate hypersensitivity, and for immunomodulation.
 CC Sequence 417 AA:

Query Match 72.7%; Score 48; DB 26; Length 417;
 Best Local Similarity 70.0%; Pred. No. 1.35e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 210 sylstsssl 219
 QY 2 sylstsssl 11

```

RESULT 14
ID W40054 standard; Protein; 481 AA.
AC W40054;
DT 20-JUL-1998 (first entry)
DE p300/CBP-associated transcriptional cofactor P/CAF C-terminus.
KW P/CAF; human; p300; CBP; transcription; cofactor;
KM histone acetyltransferase; HIV; infection; cancer; therapy;
KW muscle differentiation.
OS Homo sapiens.
PN W09803652-A2.
PD 29-JAN-1998.
PE 23-JUL-1997; U12877.
PR 23-JUL-1996; US-022273.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Howard BH, Nakatani Y;
DR MPI; 98-120777/11.
PT New isolated p300/CBP-associated factor, P/CAF - used to develop
products for modulating transcription, e.g. for treating HIV
infection or cancers or for promoting muscle differentiation
Claim 2: Page 72-73: 107pp: English.
CC This polypeptide comprises the C-terminal region (amino acid
residues 352-832) of a novel human p300/CBP associated cofactor,
P/CAF (see W40052), that modulates transcription through binding
to the cell transcription cofactors p300 (see W40055) and CBP
(see W40060) and through acetylation of histones. This C-terminal
region contains the histone acetyltransferase activity of the
protein. The invention provides methods of screening for compounds
that inhibit or stimulate the transcription modulating and histone
acetyltransferase activity of P/CAF and p300/CBP. Inhibitors can
be used e.g. to inhibit HIV TAR-mediated transcription in the
treatment of HIV infection. Stimulators can be used e.g. to
activate tumour suppressor p53 in the treatment of cancer or to
activate the muscle differentiation factor MyoD to promote muscle
differentiation. The products can also be used to inhibit the cell
cycle progression inducing effect of an oncoprotein which binds
p300/CBP in a subject. The C-terminal fragment is specific for
P/CAF and can be used to identify and define P/CAF.
SQ Sequence 481 AA:

Query Match 72.7%; Score 48; DB 30; Length 481;
Best Local Similarity 63.6%; Pred. No. 1.35e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 43 isynstsssl 53
:111111111:
1 VSYLSTRASSLD 11

RESULT 15
ID W40052 standard; Protein; 832 AA.
AC W40052;
DT 20-JUL-1998 (first entry)
DE Human p300/CBP-associated transcriptional cofactor P/CAF.
KW P/CAF; human; p300; CBP; transcription; cofactor;
KM histone acetyltransferase; HIV; infection; cancer; therapy;
KW muscle differentiation.
OS Homo sapiens.
PN W09803652-A2.
PD 29-JAN-1998.
PE 23-JUL-1997; U12877.
PR 23-JUL-1996; US-022273.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Howard BH, Nakatani Y;
DR MPI; 98-120777/11.

```

```

PT New isolated p300/CBP-associated factor, P/CAF - used to develop
products for modulating transcription, e.g. for treating HIV
infection or cancers or for promoting muscle differentiation
PS Claim 1: Page 71-72: 107pp: English.
CC This protein is a novel human p300/CBP associated cofactor,
designated P/CAF, that modulates transcription through binding to
the cell transcription cofactors p300 (see W40055) and CBP (see
W40060) and through acetylation of histones. Its amino acid
sequence was deduced from a clone (see V10090) obtained from a
human foetal liver cDNA library. The N-terminal region (see
W40056) of P/CAF, which contains the binding site for p300/CBP,
and the C-terminal region (see W40054), which contains the histone
acetyltransferase activity are also claimed. The invention provides
methods of screening for compounds that inhibit or stimulate the
transcription modulating and histone acetyltransferase activity of
P/CAF and p300/CBP. Inhibitors can be used e.g. to inhibit HIV
TAR-mediated transcription in the treatment of HIV infection.
Stimulators can be used e.g. to activate tumour suppressor p53 in
the treatment of cancer or to activate the muscle differentiation
factor MyoD to promote muscle differentiation. The products can
also be used to inhibit the cell cycle progression inducing effect
of an oncoprotein which binds p300/CBP in a subject. The products
can also be used for detection, screening assays and production of
transgenic animals.
SQ Sequence 832 AA:

Query Match 72.7%; Score 48; DB 30; Length 832;
Best Local Similarity 63.6%; Pred. No. 1.35e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 394 isynstsssl 404
:111111111:
1 VSYLSTRASSLD 11

Search completed: Thu Sep 2 11:20:37 1999
Job time : 19 secs.

```

THIS PAGE BLANK (USPTO)



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msearch_pp protein - protein database search, using Smith-Waterman algorithm

on: Thu Sep 2 12:30:26 1999; Maspar time 3.46 Seconds

73.740 Million cell updates/sec

Output not generated.

Title: >US-08-599-226-29

Description: (1-12) from US08599226.pep

Perfect Score: 75

Sequence: 1 ASYLSTSSLDY 12

Scoring table: PAM 150

Gap 15

Searched: 170751 seqs, 2126608 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

a-geneeq35
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39

Statistics: Mean 17.262; Variance 57.060; scale 0.303

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	75	100.0	12 27	W27588	Anti-TNF-alpha antilo	4.55e-01
2	72	96.0	12 27	W27593	Anti-TNF-alpha antilo	9.58e-01
3	69	92.0	12 27	W27592	Anti-TNF-alpha antilo	2.00e+00
4	69	92.0	12 27	W27591	Anti-TNF-alpha antilo	2.00e+00
5	69	92.0	12 27	W27590	Anti-TNF-alpha antilo	2.00e+00
6	67	89.3	12 27	W27589	Anti-TNF-alpha antilo	3.28e+00
7	63	84.0	12 27	W27588	Anti-TNF-alpha antilo	8.55e+00
8	63	84.0	12 27	W27589	Anti-TNF-alpha antilo	8.55e+00
9	63	84.0	12 27	W27590	Anti-TNF-alpha antilo	8.55e+00
10	57	76.0	12 27	W27593	Anti-TNF-alpha antilo	3.53e+01
11	57	76.0	12 27	W27592	Anti-TNF-alpha antilo	3.53e+01
12	54	72.0	159 37	W80222	Human succinate-ubiqu	7.08e+01
13	54	72.0	159 37	W85260	H. pylori ORF 06cp306	7.08e+01
14	54	72.0	177 29	W55472	H. pylori ORF 06cp306	7.08e+01
15	54	72.0	526 29	W55688	H. pylori ORF 09cp107	7.08e+01
16	54	72.0	875 28	W34987	Bankia gouldi endoglu	7.08e+01

17	50	66.7	166 32	W59345	Human ErbB2 domain 1	1.76e+02
18	50	66.7	191 32	W59354	Human ErbB2 protein f	1.76e+02
19	50	66.7	191 32	W23067	Canine IgE heavy chai	1.76e+02
20	50	66.7	782 23	W19764	Her2-GM-CSF immuno	1.76e+02
21	50	66.7	1255 19	W01111	HER-2/neu protein.	1.76e+02
22	50	66.7	1433 8	R39568	Sequence of c-erbB-2	1.76e+02
23	49	65.3	37 22	W17907	Peptide CS4 from dena	2.20e+02
24	49	65.3	37 26	W24223	Peptide fragment from	2.20e+02
25	49	65.3	117 22	W17913	Peptide CS4 from dena	2.20e+02
26	49	65.3	384 20	W06968	M. tuberculosis ORF 2	2.20e+02
27	49	65.3	409 34	W11495	Helicobacter polypept	2.20e+02
28	49	65.3	624 2	R08222	Extracellular portion	2.20e+02
29	49	65.3	1026 11	R58906	Human protocadherin-4	2.20e+02
30	49	65.3	1026 17	R87146	Protocadherin clone 4	2.20e+02
31	49	65.3	1203 17	R87152	Alternatively spliced	2.20e+02
32	49	65.3	1203 11	R58911	Product of alternativ	2.20e+02
33	48	64.0	459 27	W36485	Mouse TUB form 1.	2.74e+02
34	48	64.0	460 27	W36488	Human TUB form 6.	2.74e+02
35	48	64.0	505 24	W10728	Mouse tub gene produc	2.74e+02
36	48	64.0	505 27	W36486	Mouse TUB form II.	2.74e+02
37	48	64.0	505 32	W54367	Mouse tub polypeptide	2.74e+02
38	48	64.0	506 24	W10729	Human tub gene produc	2.74e+02
39	48	64.0	506 27	W36497	Human TUB form 4.	2.74e+02
40	48	64.0	506 32	W54368	Human tub polypeptide	2.74e+02
41	48	64.0	512 27	W36496	Human TUB form 3.	2.74e+02
42	48	64.0	518 27	W36495	Human TUB form 2.	2.74e+02
43	48	64.0	561 27	W36489	Human TUB form 1.	2.74e+02
44	48	64.0	2799 39	W81867	Human tumour suppress	2.74e+02
45	47	62.7	290 18	R97837	Kaposi's sarcoma asso	3.42e+02

ALIGNMENTS

RESULT 1
ID W27588 standard; peptide: 12 AA.
AC W27588;
DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HIVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-AI.
PD 14-AUG-1997.
PE 10-FEB-1997; U02219.
PR 25-NOV-1996; US-031476.
PR 09-FEB-1996; US-599226. *msf*
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,
PI Manukovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfield JG, Schoenhaut D, Vaughan TV, White M, Willton AJ;
PI WPI; 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity; e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 73; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple

CC sclerodermis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF- α induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUYEC).
Sequence 12 AA:

Query Match	100.0%;	Score 75;	DB 27;	length 12;
Best Local Similarity	100.0%;	Pred. No. 4.55e-01;		
Matches	12;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

```

Db      1 asy1stssldy 12
        |||||
QY      1 ASYLSTSSSLDY 12

```

SULT 2
W27593 standard; peptide; 12 AA.
W27593.
AC

DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
PD 14-AUG-1997.
PE 10-FEB-1997; U022219.
PR 25-NOV-1996; US-031476.
PR 09-FEB-1996; US-599226.
PA (BAD1) BASE AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakçalan Z, Labkovsky B,
PI Matkovich JG, McGuiness BT, Roberts AJ, Sforzafes P,
PI Salfeld JG, Scheemnaat D, Vaughan TJ, White M, Wilton AJ;
PR WPI: 97-413502/38.

Query Match	96.08;	Score 72;	DB 27;	Length 12;
Best Local Similarity	83.38;	Pred. No. 9.58e-01;		

	Matches	10; Conservative	2; Mismatches	0; Indels	0; Gaps	0;
Db	1	asylstassley	12			
		: :				
Qy	1	ASyLSTSSLDY	12			

RESULT	3
ID	W27592 standard; peptide; 12 AA
AC	W27592

DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HIVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN MO9729131-A1.
PD 14-AUG-1997.
PR 10-FEB-1997; U02219.
PR 25-NOV-1996; US-031476.
PR 09-FEB-1996; US-599226.

Query Match	92.0%;	Score 69;	DB 27;	Length 12;
Best Local Similarity	83.3%;	Pred. No. 2.00e+00;		
Matches	10;	Conservative	2;	Mismatches 0;
			Indels	0;
			Gaps	0;

```

Db      1 asflstsssley 12
        ||:|||||:|
QY      1 ASYLSTSSSLDY 12

```

RESULT	4
ID	W27591
	standard; peptide; 12 AA

AC W2/591, (first entry)
 DT 19-MAR-1998
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3

KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; U02219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
 WPI: 97-415302/38.
 High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:
 SQ
 Query Match 92.0%; Score 69; DB 27; Length 12;
 Best Local Similarity 91.7%; Pred. No. 2.00e+00;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 asylstssslhy 12
 1 ASYLSTSSSLDY 12
 OY
 RESULT 5
 ID W27569 standard; protein; 121 AA.
 AC W27569:
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain variable region.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody;
 KW heavy chain; variable region; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; U02219.

PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
 WPI: 97-415302/38.
 DR N-PSDB; T88404.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 16; Page 76; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain variable region.
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 121 AA:
 SQ
 Query Match 92.0%; Score 69; DB 27; Length 121;
 Best Local Similarity 90.9%; Pred. No. 2.00e+00;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 100 sylvstssslgy 110
 2 SYLSTSSSLDY 12
 OY
 RESULT 6
 ID W27590 standard; peptide; 12 AA.
 AC W27590:
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; U02219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuinness BR, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
 WPI: 97-415302/38.
 DR High affinity antibodies against human TNF alpha - useful to inhibit
 DR TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

Query Match 89.3%; Score 67; DB 27; Length 12;
Best Local Similarity 91.7%; Pred. No. 3,26e+00;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 asylstsfslidy 12
1 asylstsfslidy 12
Oy 1 ASYLSTSSSLDY 12

RESULT 7
ID W27586 standard; peptide: 12 AA.
AC W27586;

DE 19-MAR-1998 (first entry)
PI Anti-TNF-alpha antibody heavy chain CDR3.

KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.

OS Homo sapiens.

PN WO9729131-A1.

PD 14-AUG-1997.

PE 10-FEB-1997; US-02219.

PR 25-NOV-1996; US-031476.

PS 09-FEB-1996; US-599226.

PI (BADI) BASF AG.

PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,

PI Markovitch JA, McGuiness BT, Roberts AJ, Sakorafas P,

PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;

PI WPI: 97-415302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer

PS Claim 20; Page 72; 102pp; English.

CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and

CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which

CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid

CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic

CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,

CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,

CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).

CC Sequence 12 AA;

Query Match 84.0%; Score 63; DB 27; Length 12;
Best Local Similarity 100.0%; Pred. No. 8,55e+00;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 asylstsfslidy 11
1 asylstsfslidy 11
Oy 1 ASYLSTSSSLD 11

CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

Query Match 84.0%; Score 63; DB 27; Length 12;
Best Local Similarity 100.0%; Pred. No. 8,55e+00;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 asylstsfslidy 11
1 asylstsfslidy 11
Oy 1 ASYLSTSSSLD 11

RESULT 8
ID W27589 standard; peptide: 12 AA.
AC W27589;

DE 19-MAR-1998 (first entry)
PI Anti-TNF-alpha antibody heavy chain CDR3.

KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.

OS Homo sapiens.

PN WO9729131-A1.

PD 14-AUG-1997.

PE 10-FEB-1997; US-02219.

PR 25-NOV-1996; US-031476.

PS 09-FEB-1996; US-599226.

PI (BADI) BASF AG.

PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,

PI Markovitch JA, McGuiness BT, Roberts AJ, Sakorafas P,

PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;

PI WPI: 97-415302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer

PS Claim 20; Page 73; 102pp; English.

CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or

CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro

CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which

CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid

CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic

CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,

CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,

CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell

CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).

CC Sequence 12 AA;

Query Match 84.0%; Score 63; DB 27; Length 12;
Best Local Similarity 100.0%; Pred. No. 8,55e+00;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 asylstsfslidy 11
1 asylstsfslidy 11
Oy 1 ASYLSTSSSLD 11

RESULT 9
ID W2587 standard; peptide: 12 AA.
AC W2587:
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW peridontal disease; obesity; radiation formation; pyrexia; HUVEC;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-AL.
14-AUG-1997.
10-FEB-1997: U02219.
25-NOV-1996: US-031476.
09-FEB-1996: US-599226.
PR (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankevich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 73: 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L29 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
CC Spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption diseases,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, peridontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SC Sequence 12 AA:
Query Match 84.0%; Score 63; DB 27; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.55e+00;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11
ID W2594 standard; peptide: 12 AA.
AC W2594:
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW peridontal disease; obesity; radiation formation; pyrexia; HUVEC;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-AL.
14-AUG-1997.
10-FEB-1997: U02219.
25-NOV-1996: US-031476.
09-FEB-1996: US-599226.
PR (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankevich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
PS Claim 9: Page 65: 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L29 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
CC Spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, peridontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SC Sequence 12 AA:
Query Match 76.0%; Score 57; DB 27; Length 12;
Best Local Similarity 90.0%; Pred. No. 3.55e+01;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HKJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McQuinn BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR WPI: 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Disclosure: Page 75; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 12 AA:

Query Match 76.0%; Score 57; DB 27; Length 12;
 Best Local Similarity 90.0%; Pred. No. 3.53e+01;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 2 ylstassld 11
 ||||:||||
 Qy 2 YLSTSSSLD 11

RESULT 12
 ID W80222 standard; Protein: 159 AA.
 AC W80222:
 DT 06-JAN-1999 (first entry)
 DE Human succinate-ubiquinone reductase membrane anchor subunit.
 KW Human; succinate-ubiquinone reductase membrane anchor subunit; SDHMA;
 KW myopathy; progressive externalphthalmoplegia; Kearns-Sayre syndrome;
 KW myoclonic epilepsy; ophthalmopathy; cardiomyopathy; lactic acidosis.
 NS Homo sapiens.
 US582711.1-A.
 27-OCT-1998.
 20-MAR-1997; 828832.
 PA (INCY-) INCYTE PHARM INC.
 PI Lal P, Shah P;
 DR WPI: 98-593998/50.
 DR N-PSDB: V66406.
 PT DNA encoding succinate-ubiquinone reductase membrane anchor subunit
 PT - useful for producing recombinant polypeptide
 PS Claim 1: Fig 1A-B; 25pp; English.
 CC The present sequence represents a human succinate-ubiquinone reductase
 CC membrane anchor subunit (SDHMA). The SDHMA nucleic acid sequence
 CC was first identified in incyte clone 2454416 from the aortic
 CC endothelial cell cDNA library. The protein has chemical and structural
 CC homology with membrane anchoring subunits from bovine mitochondria,
 CC OPA3 and OPS1. The protein also contains potential intramolecular
 CC disulphide bridging sites, which are found at Cys11, Cys44, Cys88 and
 CC Cys150. SDHMA can be used to treat myopathies, e.g. progressive
 CC externalphthalmoplegia, Kearns-Sayre syndrome, myoclonic epilepsy,
 CC cardiomyopathy or lactic acidosis.
 SQ Sequence 159 AA:

Query Match 72.0%; Score 54; DB 37; Length 159;
 Best Local Similarity 50.0%; Pred. No. 7.08e+01;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 Db 82 aaylncsamdy 93
 |::|:|:|
 Qy 1 ASYLSTSSSLDY 12

RESULT 13
 ID W55260 standard; Protein: 159 AA.
 AC W55260:
 DT 02-JUL-1998 (first entry)
 DE H. pylori ORF 06cp0603orf16 protein.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW identification; binding compound; bacteria; life cycle; activator;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 OS Helicobacter pylori.
 PN W09737044-A1.
 PD 09-OCT-1997.
 PF 27-MAR-1997; U05223.
 PR 06-DEC-1996; US-761318.
 PR 29-MAR-1996; US-625811.
 PR 02-APR-1996; US-758731.
 PR 25-OCT-1996; US-736905.
 PR 28-OCT-1996; US-738859.
 PA (ASTR) ASTRA AB.
 PI Alm RA, Smith D;
 DR WPI: 97-503122/46.
 DR N-PSDB: V24669.
 PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection
 PS Claim 14; Page 502; 1145pp; English.
 CC This sequence is a H. pylori protein of unspecified function.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance
 CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 51679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. To identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 SQ Sequence 159 AA:

Query Match 72.0%; Score 54; DB 29; Length 159;
 Best Local Similarity 60.0%; Pred. No. 7.08e+01;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 60 ylstasaley 69
 |||:|:|:|
 Qy 3 YLSTSSSLDY 12

RESULT 14
 ID W55472 standard; Protein: 177 AA.
 AC W55472:
 DT 24-JUN-1998 (first entry)
 DE H. pylori ORF 06cp0603_23452_c3.80 inner membrane protein.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW identification; binding compound; bacteria; life cycle; activator;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 OS Helicobacter pylori.
 PN W09737044-A1.

PD 09-OCT-1997. 005223.
 PF 27-MAR-1997. 005223.
 PR 06-DEC-1996; US-761318.
 PR 29-MAR-1996; US-625811.
 PR 02-APR-1996; US-758731.
 PR 25-OCT-1996; US-736905.
 PR 28-OCT-1996; US-738859.
 PA (ASTR) ASTRA AB.
 PI Alm RA, Smith D;
 DR WPI: 97-503122/46.
 DR N-PSDB: V24881.
 PT Helicobacter pylori nucleic acid sequences and encoded polypeptide(s) - useful in vaccines to treat or prevent H. pylori infection and for diagnosis of H. pylori infection
 PS Claims 14,80; Page 679; 1145pp; English.
 CC This sequence is a H. pylori cell envelope inner membrane protein involved in cofactor metabolism.
 CC The protein may be used in a vaccine to prevent or treat H. pylori infection or to identify H. pylori polypeptide binding compounds, useful as potential H. pylori life cycle activators or inhibitors. The DNA and probes derived from it may be used for the identification of H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic acid sequences complementary to the DNA act as antisense sequences and can be used to prevent the translation of H. pylori mRNA. Antibodies against the protein can be used in immunoassays to evaluate the abundance and distribution of H. pylori-specific antigens. The genomic sequence of H. pylori (ATCC 55679) was determined from overlapping contigs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 180 nucleotides, and the predicted coding regions defined by computer evaluation. To identify likely H. pylori antigens for vaccine development, the amino acid sequences predicted from various ORF were analysed for significant homology to other known or exported membrane proteins. Having identified and determined the sequences of interest, particular regions can be isolated from H. pylori by PCR amplification for recombinant polypeptide production, e.g. in E. coli hosts.
 CC Sequence 177 AA;

Query Match 72.0%; Score 54; DB 29; Length 177;
 Best Local Similarity 60.0%; Pred. No. 7.08e+01;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 60 ylsalasaley 69
 |||:|:|:|
 QY 3 YLSTSSLDY 12

RESULT 15
 W55688 standard; Protein; 526 AA.
 W55688;
 DT 07-JUL-1998 (first entry)
 DE H. pylori ORF 09cpl0713_23452_c3_195 inner membrane protein.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope; identification; binding compound; bacteria; life cycle; activator;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 OS Helicobacter pylori.
 PN MO9737064-A1.
 PD 09-OCT-1997.
 PF 27-MAR-1997; 005223.
 PR 06-DEC-1996; US-761318.
 PR 29-MAR-1996; US-625811.
 PR 02-APR-1996; US-758731.
 PR 25-OCT-1996; US-736905.
 PR 28-OCT-1996; US-738859.
 PA (ASTR) ASTRA AB.
 PI Alm RA, Smith D;
 DR WPI: 97-503122/46.
 DR N-PSDB: V25097.
 PT Helicobacter pylori nucleic acid sequences and encoded polypeptide(s) - useful in vaccines to treat or prevent H. pylori infection and for diagnosis of H. pylori infection
 PS Claims 14,80; Pages 947-948; 1145pp; English.
 CC This sequence is a H. pylori cell envelope inner membrane protein

CC Involved in cofactor metabolism.
 CC The protein may be used in a vaccine to prevent or treat H. pylori infection or to identify H. pylori polypeptide binding compounds, useful as potential H. pylori life cycle activators or inhibitors. The DNA and probes derived from it may be used for the identification of H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic acid sequences complementary to the DNA act as antisense sequences and can be used to prevent the translation of H. pylori mRNA. Antibodies against the protein can be used in immunoassays to evaluate the abundance and distribution of H. pylori-specific antigens. The genomic sequence of H. pylori (ATCC 55679) was determined from overlapping contigs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 180 nucleotides, and the predicted coding regions defined by computer evaluation. To identify likely H. pylori antigens for vaccine development, the amino acid sequences predicted from various ORF were analysed for significant homology to other known or exported membrane proteins. Having identified and determined the sequences of interest, particular regions can be isolated from H. pylori by PCR amplification for recombinant polypeptide production, e.g. in E. coli hosts.
 CC Sequence 526 AA;

Query Match 72.0%; Score 54; DB 29; Length 526;
 Best Local Similarity 60.0%; Pred. No. 7.08e+01;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 60 ylsalasaley 69
 |||:|:|:|
 QY 3 YLSTSSLDY 12

Search completed: Thu Sep 2 12:30:47 1999
 Job time : 21 secs.

THIS PAGE BLANK (USPTO)

WORLDWIDE
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSrch.p protein - protein database search, using Smith-Waterman algorithm

on: Thu Sep 2 12:33:13 1999; Maspar time 1.39 Seconds
87.360 Million cell updates/sec

Modular output not generated.

Title: >US-08-599-226-29
Description: (1-12) from US08599226.pep
Perfect Score: 75
Sequence: 1 ASYLSTSSSLDY 12

Scoring table: PAM 150
Gap 15

Searched: 106580 seqs, 10152877 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PC9_COMB 4:backfiles1

Statistics: Mean 16.112; Variance 54.206; scale 0.297

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	54	72.0	158	2	US-08-828-Sequence 3, Applicatio	3.02e+01
2	54	72.0	159	2	US-08-828-Sequence 1, Applicatio	3.02e+01
3	50	66.7	426	1	PCT-US95-1Sequence 2, Applicatio	7.56e+01
4	50	66.7	426	1	US-08-336-Sequence 2, Applicatio	7.56e+01
5	50	66.7	1255	2	US-08-414-Sequence 68, Applicati	7.56e+01
6	50	66.7	1255	1	US-08-467-Sequence 68, Applicati	7.56e+01
7	50	66.7	1255	2	US-08-625-Sequence 2, Applicatio	7.56e+01
8	50	66.7	1255	2	US-08-486-Sequence 68, Applicati	7.56e+01
9	50	66.7	1255	2	US-08-484-Sequence 8, Applicatio	7.56e+01
10	50	66.7	1255	2	US-08-356-Sequence 68, Applicati	7.56e+01
11	50	66.7	1255	2	US-08-356-Sequence 2, Applicatio	7.56e+01
12	49	65.3	37	3	PCT-US96-0Sequence 5, Applicatio	9.48e+01
13	49	65.3	117	3	PCT-US96-0Sequence 11, Applicati	9.48e+01
14	49	65.3	384	1	US-08-457-Sequence 5, Applicatio	9.48e+01
15	49	65.3	1026	1	PCT-US95-0Sequence 95, Applicati	9.48e+01
16	49	65.3	1026	1	US-08-453-Sequence 95, Applicati	9.48e+01
17	49	65.3	1026	1	US-08-453-Sequence 95, Applicati	9.48e+01
18	49	65.3	1026	1	US-07-998-Sequence 95, Applicati	9.48e+01
19	49	65.3	1026	2	US-08-268-Sequence 95, Applicati	9.48e+01
20	49	65.3	1026	3	PCT-US93-1Sequence 95, Applicati	9.48e+01
21	49	65.3	1203	3	PCT-US95-0Sequence 103, Applicat	9.48e+01
22	49	65.3	1203	3	PCT-US93-1Sequence 103, Applicat	9.48e+01
23	49	65.3	1203	2	US-08-268-Sequence 103, Applicat	9.48e+01

24	49	65.3	1203	1	US-08-453-Sequence 103, Applicat	9.48e+01
25	49	65.3	1203	1	US-08-453-Sequence 103, Applicat	9.48e+01
26	49	65.3	1203	1	US-07-998-Sequence 103, Applicat	9.48e+01
27	48	64.0	459	2	US-08-630-Sequence 2, Applicatio	1.19e+02
28	48	64.0	459	2	US-08-714-Sequence 2, Applicatio	1.19e+02
29	48	64.0	460	2	US-08-714-Sequence 2, Applicatio	1.19e+02
30	48	64.0	460	2	US-08-630-Sequence 7, Applicatio	1.19e+02
31	48	64.0	505	2	US-08-936-Sequence 2, Applicatio	1.19e+02
32	48	64.0	505	2	US-08-630-Sequence 4, Applicatio	1.19e+02
33	48	64.0	505	2	US-08-936-Sequence 2, Applicatio	1.19e+02
34	48	64.0	505	2	US-08-714-Sequence 4, Applicatio	1.19e+02
35	48	64.0	505	2	US-08-922-Sequence 2, Applicatio	1.19e+02
36	48	64.0	505	2	US-08-829-Sequence 2, Applicatio	1.19e+02
37	48	64.0	505	1	US-08-631-Sequence 2, Applicatio	1.19e+02
38	48	64.0	506	2	US-08-936-Sequence 8, Applicatio	1.19e+02
39	48	64.0	506	2	US-08-829-Sequence 8, Applicatio	1.19e+02
40	48	64.0	506	2	US-08-922-Sequence 8, Applicatio	1.19e+02
41	48	64.0	506	2	US-08-936-Sequence 8, Applicatio	1.19e+02
42	48	64.0	506	1	US-08-631-Sequence 8, Applicatio	1.19e+02
43	48	64.0	561	2	US-08-714-Sequence 27, Applicati	1.19e+02
44	47	62.7	290	2	US-08-420-Sequence 27, Applicati	1.49e+02
45	47	62.7	290	3	PCT-US95-1Sequence 27, Applicati	1.49e+02

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	158 AA.
ID	US-08-828-832-3			
XX	xxxxxx			
AC				
XX				
DT				
XX				
DE	Sequence 3, Application US/08828832			
XX				
CC	Sequence 3, Application US/08828832			
CC	Patent No. 5827711			
CC	GENERAL INFORMATION:			
CC	APPLICANT: Lal, Preeti			
CC	APPLICANT: Shah, Puri			
CC	TITLE OF INVENTION: NOVEL SUCCINATE DEHYDROGENASE SUBUNIT			
CC	NUMBER OF SEQUENCES: 4			
CC	CORRESPONDENCE ADDRESS:			
CC	ADDRESSEE: Incyte Pharmaceuticals, Inc.			
CC	STREET: 3174 Porter Drive			
CC	CITY: Palo Alto			
CC	STATE: CA			
CC	COUNTRY: USA			
CC	ZIP: 94304			
CC	COMPUTER READABLE FORM:			
CC	MEDIUM TYPE: Diskette			
CC	COMPUTER: IBM Compatible			
CC	OPERATING SYSTEM: DOS			
CC	SOFTWARE: FastSeq for Windows Version 2.0			
CC	CURRENT APPLICATION DATA:			
CC	APPLICATION NUMBER: US/08/828,832			
CC	FILING DATE: Herewith			
CC	CLASSIFICATION: 435			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER:			
CC	FILING DATE:			
CC	ATTORNEY/AGENT INFORMATION:			
CC	NAME: Billings, Lucy J.			
CC	REGISTRATION NUMBER: 36,749			
CC	REFERENCE/DOCKET NUMBER: PF-0250 US			
CC	TELECOMMUNICATION INFORMATION:			
CC	TELEPHONE: 415-855-0555			
CC	TELEFAX: 415-845-4166			
CC	TELEX:			
CC	INFORMATION FOR SEQ ID NO: 3:			
CC	SEQUENCE CHARACTERISTICS:			
CC	LENGTH: 158 amino acids			
CC	TYPE: amino acid			

CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: GenBank
CC CLONE: 1575011
SQ SEQUENCE 158 AA: 17096 MW: 135076 CN:

Query Match 72.0%; Score 54; DB 2; Length 158;
Best Local Similarity 50.0%; Pred. No. 3.02e+01;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 81 AAYLNPCSAMLY 92
1:11: 1:11:
QY 1 ASYLSTSSLDY 12

RESULT 2
ID US-08-828-832-1 STANDARD; PRT: 159 AA.

XX xxxxxx

DE Sequence 1, Application US/08828832

CC Sequence 1, Application US/08828832

CC Patent No. 5827711

CC GENERAL INFORMATION:

CC APPLICANT: Lal, Preeti

CC APPLICANT: Shah, Purvi

CC TITLE OF INVENTION: NOVEL SUCCINATE DEHYDROGENASE SUBUNIT

CC NUMBER OF SEQUENCES: 4

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Incyte Pharmaceuticals, Inc.

CC STREET: 3174 Porter Drive

CC CITY: Palo Alto

CC STATE: CA

CC COUNTRY: USA

CC ZIP: 94304

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Diskette

CC COMPUTER: IBM Compatible

CC OPERATING SYSTEM: DOS

CC SOFTWARE: FastSeq for Windows Version 2.0

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/828, 832

CC FILING DATE: Herewith

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER:

CC FILING DATE:

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Billings, Lucy J

CC REGISTRATION NUMBER: 36,749

CC REFERENCE/DOCKET NUMBER: PF-0250 US

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 415-855-0555

CC TELEFAX: 415-845-4166

CC TELEX:

CC INFORMATION FOR SEQ ID NO: 1:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 159 amino acids

CC TYPE: amino acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC IMMEDIATE SOURCE:

CC LIBRARY: Consensus

CC CLONE: 2454416

SQ SEQUENCE 159 AA: 17043 MW: 129477 CN;

Query Match 72.0%; Score 54; DB 2; Length 159;

Best Local Similarity 50.0%; Pred. No. 3.02e+01;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 82 AAYLNPCSAMLY 93
1:11: 1:11:
QY 1 ASYLSTSSLDY 12

RESULT 3
ID PCT-US95-13795-2 STANDARD; PRT: 426 AA.

XX xxxxxx

DE Sequence 2, Application PC/TUS9513795

CC Sequence 2, Application PC/TUS9513795

CC GENERAL INFORMATION:

CC APPLICANT: HOLDIS, GREGORY F.

CC APPLICANT: PATEL, MAYUR D.

CC TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULINS

CC NUMBER OF SEQUENCES: 4

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CHRISTINE E. CARTY

CC STREET: 126 E. LINCOLN AVENUE; P.O. BOX 2000

CC CITY: RAHWAY

CC STATE: NEW JERSEY

CC COUNTRY: USA

CC ZIP: 07065-0907

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US95/13795

CC FILING DATE:

CC CLASSIFICATION:

CC ATTORNEY/AGENT INFORMATION:

CC NAME: CARTY, CHRISTINE E.

CC REGISTRATION NUMBER: 36,099

CC REFERENCE/DOCKET NUMBER: 19211Y

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (908) 594-6734

CC TELEFAX: (908) 594-4720

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 426 amino acids

CC TYPE: amino acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: Protein

SQ SEQUENCE 426 AA: 47234 MW: 1032622 CN;

Query Match 66.7%; Score 50; DB 3; Length 426;
Best Local Similarity 63.6%; Pred. No. 7.56e+01;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 214 TSYLSPSPPLD 224
1:111: 1:111:
QY 1 ASYLSTSSLDY 11

RESULT 4
ID US-08-336-583-2 STANDARD; PRT: 426 AA.

XX xxxxxx

DE Sequence 2, Application US/08336583

CC Sequence 2, Application US/08336583

CC Patent No. 5629415

CC GENERAL INFORMATION:
CC APPLICANT: HOLDS, GREGORY F.
CC APPLICANT: PATEL, MAYUR D.
CC TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN E
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CHRISTINE E. CARTY
CC STREET: 126 E. LINCOLN AVENUE
CC CITY: RAHWAY
CC STATE: NEW JERSEY
CC COUNTRY: USA
CC ZIP: 07065-0900
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/336,583
CC FILING DATE: 09-NOV-1994
CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: CARTY, CHRISTINE E.
CC REGISTRATION NUMBER: 36,099
CC REFERENCE/DOCKET NUMBER: 19211
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (908) 594-6734
CC TELEFAX: (908) 594-4720
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 426 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 426 AA; 47234 MW; 1032622 CN;
SQ

Query Match 66.7%; Score 50; DB 1; Length 426;
Best Local Similarity 63.6%; Pred. No. 7.56e+01;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 214 TSYLSPSPSLD 224
:||||:|:
0y 1 ASYLSTSSSLD 11

RESULT 5 STANDARD: PRT: 1255 AA.
DB US-08-414-417B-68
xxxxxx

Sequence 68, Application US/08414417B
Patent No. 5801005
GENERAL INFORMATION:
CC APPLICANT: Cheever, Martin A.
CC APPLICANT: Disis, Mary L.
CC TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
CC TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
CC TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
CC NUMBER OF SEQUENCES: 69
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Seed and Berry LLP
CC STREET: 6300 Columbia Center, 701 Fifth Avenue
CC CITY: Seattle
CC STATE: Washington
CC COUNTRY: US
CC ZIP: 98104-7092
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/414,417B
CC FILING DATE: 31-MAR-1995
CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sharkey, Richard G.
CC REGISTRATION NUMBER: 32,629
CC REFERENCE/DOCKET NUMBER: 920010.448C2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 622-4900
CC TELEFAX: (206) 682-6031
CC INFORMATION FOR SEQ ID NO: 68:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1255 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC SEQUENCE 1255 AA; 137955 MW; 8109851 CN;
SQ

Query Match 66.7%; Score 50; DB 2; Length 1255;
Best Local Similarity 45.5%; Pred. No. 7.56e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Db 63 TYLPTNASTLSF 73
:||||:|:
0y 2 SYLSTSSSLD 12

RESULT 6 STANDARD: PRT: 1255 AA.
DB US-08-467-083-68
xxxxxx

Sequence 68, Application US/08467083
Patent No. 5726023
GENERAL INFORMATION:
CC APPLICANT: Cheever, Martin A.
CC APPLICANT: Disis, Mary L.
CC TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
CC TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH TH
CC NUMBER OF SEQUENCES: 68
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Seed and Berry
CC STREET: 6300 Columbia Center, 701 Fifth Avenue
CC CITY: Seattle
CC STATE: Washington
CC COUNTRY: US
CC ZIP: 98104-7092
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467,083
CC FILING DATE: 06-JUN-1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/414,417
CC FILING DATE: 06-JUN-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sharkey, Richard G.
CC REGISTRATION NUMBER: 32,629
CC REFERENCE/DOCKET NUMBER: 920010.448C2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 622-4900

CC TELEFAX: (206) 682-6031
CC TELEX: 3723836 SEDANBERRY
CC INFORMATION FOR SEQ ID NO: 68:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1255 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
SQ SEQUENCE 1255 AA; 137955 MW; 8109851 CN;

Query Match 66.7%; Score 50; DB 1; Length 1255;
Best Local Similarity 45.5%; Pred. No. 7.56e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 63 TYLPTNASLSF 73
QY 2 SYLSTSSSLDY 12

RESULT 7
US-08-625-101-2 STANDARD; PRT: 1255 AA.

Sequence 2, Application US/08625101
Patent No. 5869445
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/625,101
FILING DATE: 01-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 1255 AA; 137909 MW; 8111405 CN;

Query Match 66.7%; Score 50; DB 2; Length 1255;
Best Local Similarity 45.5%; Pred. No. 7.56e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 63 TYLPTNASLSF 73

QY 2 SYLSTSSSLDY 12
RESULT 8
ID US-08-486-348A-68 STANDARD; PRT: 1255 AA.
AC xxxxxx
XX
XX
XX
DE Sequence 68, Application US/08486348A
CC Sequence 68, Application US/08486348A
CC Patent No. 5846538
CC GENERAL INFORMATION:
CC APPLICANT: Cheever, Martin A.
CC APPLICANT: Disis, Mary L.
CC TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
CC TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH TH
CC TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
CC NUMBER OF SEQUENCES: 69
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Seed and Berry LLP
CC STREET: 6300 Columbia Center, 701 Fifth Avenue
CC CITY: Seattle
CC STATE: Washington
CC COUNTRY: US
CC ZIP: 98104-7092
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/486,348A
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Sharkey, Richard G.
CC REGISTRATION NUMBER: 32,629
CC REFERENCE/DOCKET NUMBER: 920010.448C6
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 622-4900
CC TELEFAX: (206) 682-6031
CC INFORMATION FOR SEQ ID NO: 68:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1255 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
SQ SEQUENCE 1255 AA; 137955 MW; 8109851 CN;

Query Match 66.7%; Score 50; DB 2; Length 1255;
Best Local Similarity 45.5%; Pred. No. 7.56e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 63 TYLPTNASLSF 73
QY 2 SYLSTSSSLDY 12

RESULT 9
ID US-08-484-438-8 STANDARD; PRT: 1255 AA.
AC xxxxxx
XX
XX
XX
DE Sequence 8, Application US/08484438
CC Sequence 8, Application US/08484438
CC Patent No. 5811098
CC Patent No. 5811098 5780031


```
CC GENERAL INFORMATION:  
CC APPLICANT: Plowman, Gregory D.  
CC APPLICANT: Culouscou, Jean-Michel  
CC APPLICANT: Shoyab, Mohammed  
CC APPLICANT: Siegal, Clay B.  
CC APPLICANT: Hellstr m, Ingegerd  
CC APPLICANT: Hellstr m, Karl E.  
CC TITLE OF INVENTION: HERA HUMAN RECEPTOR TYROSINE KINASE  
CC NUMBER OF SEQUENCES: 42  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC City: New York  
CC STATE: New York  
CC COUNTRY: U.S.A.  
CC ZIP: 10036-2711  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: floppy disk  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/484,438  
CC FILING DATE: 07-JUN-1995  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 08/333,442  
CC FILING DATE: 14-OCT-1994  
CC APPLICATION NUMBER: US 08/150,704  
CC FILING DATE: 10-NOV-1993  
CC CLASSIFICATION: 530  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 07/981,165  
CC FILING DATE: 24-NOV-1992  
CC CLASSIFICATION: 530  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Mistrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE//DOCKET NUMBER: 5624-230  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-8864/9741  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 8:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1255 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: unknown  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
CC SEQUENCE 1255 AA; 137815 MW; 8105484 CN;  
  
Cc Query Match 66.7%; Score 50; DB 2; Length 1255;  
Cc Best Local Similarity 45.5%; Pred. No. 7.5e+01;  
Cc Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
  
Cc Db 63 TYLEPTNASLSIF 73  
Cc :|::||::|  
Cc 2 SYLSTSSLDY 12  
  
Cc RESULT 10  
Cc ID US-08-468-545B-68 STANDARD: PRT; 1255 AA.  
Cc AC xxxxxx  
Cc DT  
Cc XX  
Cc DE Sequence 68, Application US/08468545B  
Cc XX  
Cc Sequence 68, Application US/08468545B  
Cc CC Patent No. 5876712  
Cc CC GENERAL INFORMATION:
```

CC	APPLICANT:	Cheever, Martin A.
CC	APPLICANT:	Disis, Mary L.
CC	TITLE OF INVENTION:	IMMUNE REACTIVITY TO HER-2/neu PROTEIN
CC	TITLE OF INVENTION:	FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH TH
CC	TITLE OF INVENTION:	HER-2/neu ONCOGENE IS ASSOCIATED
CC	NUMBER OF SEQUENCES:	69
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE:	Seed and Berry LLP
CC	STREET:	6300 Columbia Center, 701 Fifth Avenue
CC	CITY:	Seattle
CC	STATE:	Washington
CC	COUNTRY:	US
CC	ZIP:	98104-7092
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE:	Floppy disk
CC	COMPUTER:	IBM PC compatible
CC	OPERATING SYSTEM:	PC-DOS/MS-DOS
CC	SOFTWARE:	PatentIn Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER:	US/08/466,545B
CC	FILING DATE:	06-JUN-1995
CC	CLASSIFICATION:	424
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME:	Sharkey, Richard G.
CC	REGISTRATION NUMBER:	32,629
CC	REFERENCE/DOCKET NUMBER:	920010.448C5
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE:	(206) 622-4900
CC	TELEFAX:	(206) 682-6031
CC	INFORMATION FOR SEO ID NO:	68:
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH:	1255 amino acids
CC	TYPE:	amino acid
CC	TOPOLOGY:	linear
SQ	SEQUENCE	1255 AA; 137955 MW; 8109851 CN;
Dd	Query Match	66.7%; Score 50; DB 2; Length 1255;
Dd	Best Local Similarity	45.5%; Pred. No. 7.56e+01;
Dd	Matches	5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Dd	63 TYLPTNASLSF 73	
Dd	2 SYLSTSSSLDY 12	
RESULT	11	
ID	US-08-356-786-2	STANDARD; PRT; 1255 AA.
XX	xxxxxx	
DE	Sequence 2, Application US/08356786	
CC	Sequence 2, Application US/08356786	
CC	Patent No. 5877305	
CC	GENERAL INFORMATION:	
CC	APPLICANT:	Huston, James S.
CC	APPLICANT:	Oppermann, Hermann
CC	APPLICANT:	Houston, D. B.
CC	TITLE OF INVENTION:	Biosynthetic Binding Protein for Cancer
CC	TITLE OF INVENTION:	Marker
CC	NUMBER OF SEQUENCES:	16
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE:	Edmund R. Pitcher, Testa, Hurwitz, & Thibault
CC	STREET:	Exchange Place, 53 State Street
CC	CITY:	Boston
CC	STATE:	Massachusetts
CC	COUNTRY:	USA
CC	ZIP:	02109
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE:	Floppy disk

CC COMPUTER: IBM PC compatible
CC
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/356,786
CC FILING DATE:
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/831,967
CC FILING DATE: 06-FEB-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Pitcher, Edmund R.
CC REGISTRATION NUMBER: 27,829
CC REFERENCE/DOCKET NUMBER: CRP-053
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 248-7000
CC TELEFAX: (617) 248-7100
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1255 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1255 AA; 137909 MW; 8111405 CN;
SQ
Query Match 66.7%; Score 50; DB 2; Length 1255;
Best Local Similarity 45.5%; Pred. No. 7.56e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
DB 63 TYLPTNASTSF 73
OY 2 SYLSTSSSLDY 12
RESULT 12
ID PCT-US96-08730-5 STANDARD; PRT; 37 AA.
AC xxxxxx
DT
DT
DT
DT
DT
DE Sequence 5, Application PC/TUS9608730
XX
CC Sequence 5, Application PC/TUS9608730
CC GENERAL INFORMATION:
CC APPLICANT: Cassels, Frederick
CC APPLICANT: Anderson, Jeffrey
CC APPLICANT: Carter, John Mark
CC TITLE OF INVENTION: Methods of Raising Antibodies Against E.
CC TITLE OF INVENTION: Coil of the Family CSF-CFA./1
CC NUMBER OF SEQUENCES: 15
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Glenna Hendricks
CC STREET: P.O. Box 2509
CC CITY: Fairfax
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22031
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/08730
CC FILING DATE: 03-JUN-1996
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hendricks, Glenna
CC REGISTRATION NUMBER: 32,535
CC REFERENCE/DOCKET NUMBER: PCT/US96/08730
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703) 591-4470
CC TELEFAX: (703) 591-4470
CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 117 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: peptide
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FRAGMENT TYPE: Internal
SQ SEQUENCE 117 AA; 12389 MW; 76297 CN;

CC TELEFAX: (703) 591-4428
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 37 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: peptide
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FRAGMENT TYPE: Internal
SQ SEQUENCE 37 AA; 3864 MW; 7776 CN;
Query Match 65.3%; Score 49; DB 3; Length 37;
Best Local Similarity 50.0%; Pred. No. 9.48e+01;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
DB 22 GSYLPTAVELTY 33
OY 1 ASYLSTSSSLDY 12
RESULT 13
ID PCT-US96-08730-11 STANDARD; PRT; 117 AA.
AC xxxxxx
DT
DT
DT
DT
DT
DE Sequence 11, Application PC/TUS9608730
XX
CC Sequence 11, Application PC/TUS9608730
CC GENERAL INFORMATION:
CC APPLICANT: Cassels, Frederick
CC APPLICANT: Anderson, Jeffrey
CC APPLICANT: Carter, John Mark
CC TITLE OF INVENTION: Methods of Raising Antibodies Against E.
CC TITLE OF INVENTION: Coil of the Family CSF-CFA./1
CC NUMBER OF SEQUENCES: 15
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Glenna Hendricks
CC STREET: P.O. Box 2509
CC CITY: Fairfax
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22031
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/08730
CC FILING DATE: 03-JUN-1996
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hendricks, Glenna
CC REGISTRATION NUMBER: 32,535
CC REFERENCE/DOCKET NUMBER: PCT/US96/08730
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703) 591-4470
CC TELEFAX: (703) 591-4428
CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 117 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: peptide
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FRAGMENT TYPE: Internal
SQ SEQUENCE 117 AA; 12389 MW; 76297 CN;

Query Match 65.3%: Score 49; DB 3; Length 117;
Best Local Similarity 50.0%: Pred. No. 9.48e+01;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

DB 22 GSYLPTAVELTY 33
:||||:|
QY 1 ASYLSTSSSIDY 12

RESULT 14
ID US-08-457-245-5 STANDARD: PRT: 384 AA.

Sequence 5, Application US/08457245
Patent No. 5573915
GENERAL INFORMATION:
APPLICANT: BARRY III, Clifton E.
TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA INVOLVED
TITLE OF INVENTION: IN THE BIOSYNTHESIS OF CYCLOPROPANATED MYCOLIC ACIDS IN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,245
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chambers, Guy W.
REGISTRATION NUMBER: 30,617
REFERENCE/DOCKET NUMBER: 15280-216000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: ORF2 protein
SEQUENCE 384 AA: 41963 MW; 701271 CN;

Query Match 65.3%: Score 49; DB 1; Length 384;
Best Local Similarity 54.5%: Pred. No. 9.48e+01;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

DB 235 SSYLPTKAUD 245
:||||:|
QY 1 ASYLSTSSSID 11

RESULT 15
ID PCT-US95-08071-95 STANDARD: PRT: 1026 AA.
AC xxxxxx
XX
XX

Sequence 95, Application PC/TUS9508071
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
ADDRESSEE: Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08071
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12588
FILING DATE: 23 DEC 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/998,003
FILING DATE: 29 DEC 1992
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 1026 AA: 111270 MW; 5611711 CN;

Query Match 65.3%: Score 49; DB 3; Length 1026;
Best Local Similarity 50.0%: Pred. No. 9.48e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

DB 423 FLOTTPLDY 432
:||||:|
QY 3 YLSTSSSIDY 12

Search completed: Thu Sep 2 12:33:21 1999
Job time : 8 secs.

THIS PAGE BLANK (USPTO)

TITLE protein-export membrane protein - Helicobacter pylori (strain J99)
ORGANISM #formal_name Helicobacter pylori
#organism strain J99
DATE 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 05-Mar-1999
ACCESSIONS D71805
REFERENCE A71800
#authors Alm, R.A.; Ling, L.S.L.; Molt, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Noonan, B.; Guild, B.C.; deJonghe, B.L.; Carmel, G.; Tummino, P.J.; Caruso, A.; Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Voyts, G.F.; Trust, T.J.
#journal Nature (1999) 397:176-180
#title Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.
#cross-references MVID:99120557
#accession D71805
#status Preliminary
#molecule_type DNA
#residues 1-526 #label ARN
#cross-references GB:AE001567; GB:AE001439; NID:94156065; PID:94156069
#experimental_source strain J99
GENETICS
#gene secD
CLASSIFICATION #superfamily protein export membrane protein secD
SUMMARY #length 526 #molecular_weight 56796 #checksum 5813
Query Match Best Local Similarity 72.0%; Score 54; DB 2; Length 526; Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 60 YLSLASLEY 69
|||:|:|:|
QY 3 YLSTSSLDY 12
RESULT 3
ENTRY S38380 #type complete
TITLE Hrcx1 protein - California red abalone
ORGANISM #formal_name Haliotis rufescens #common_name California red abalone
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Oct-1997
ACCESSIONS S38380
REFERENCE S38380
#authors Degnan, B.M.; Degnan, S.M.; Morse, D.E.
#submission submitted to the EMBL Data Library, September 1993
#description Expression of Hrcx1, a gastropod mollusc Hox homeobox gene, is progressively restricted during morphogenesis from trochophore to veliger larval forms.
#accession S38380
#status Preliminary
#molecule_type mRNA
#residues 1-265 #label DEG
#cross-references EMBL:X75217; NID:9407414; PID:9407415
CLASSIFICATION #superfamily unassigned homeobox proteins; homeobox homology
KEYWORDS DNA binding; homeobox; nucleus; transcription regulation
FEATURE
164-220 #domain homeobox homology #label HOX
SUMMARY #length 265 #molecular_weight 29579 #checksum 8996
Query Match Best Local Similarity 70.7%; Score 53; DB 2; Length 265; Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 81 GYLSSSSSKDY 92
:|||||:|:|:|
QY 1 ASYLSTSSLDY 12
RESULT 4

ENTRY ACCE #type complete
TITLE glutamate receptor precursor - great pond snail
ORGANISM #formal_name Lymnaea stagnalis #common_name great pond snail
DATE 31-Mar-1992 #sequence_revision 28-Oct-1994 #text_change 05-Sep-1997
ACCESSIONS S18443; S15681
REFERENCE S18443
#authors Hutcheon, M.L.; Harvey, R.J.; Barnard, E.A.; Darlison, M.G.
#journal FEBS Lett. (1991) 292:111-114
#title Cloning of a cDNA that encodes an invertebrate glutamate receptor subunit.
#cross-references MVID:92070466
#accession S18443
#status nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-917 #label HUT
#cross-references EMBL:X60086
REFERENCE S15681
#authors Hutcheon, M.L.; Bhandal, N.S.; Harvey, R.J.; Usherwood, P.N.R.; Darlison, M.G.
#submission submitted to the EMBL Data Library, June 1991
#description PCR-mediated cloning of a cDNA that encodes a functional molluscan glutamate receptor subunit.
#accession S15681
#molecule_type mRNA
#residues 1-362, 'K', 364-776, 'S', 778-845, 'R', 847-886, 'S', 888-917
#label HUT2
#cross-references EMBL:X60086; NID:99628; PID:99629
CLASSIFICATION #superfamily glutamate receptor; glutamate receptor homology
KEYWORDS glycoprotein; ion channel; neurotransmitter receptor; transmembrane protein
FEATURE
1-19 #domain signal sequence #status predicted #label SIG
20-917 #product glutamate receptor #status predicted #label MAT
429-853 #domain glutamate receptor homology #label GRH
559-578 #domain transmembrane #status predicted #label TM1
599-617 #domain transmembrane #status predicted #label TM2
628-646 #domain transmembrane #status predicted #label TM3
819-839 #domain transmembrane #status predicted #label TM4
62,95,121,125,229, 251,261,272,418, 419,424,491,881 #binding_site carbohydrate (Asn) (covalent) #status predicted
SUMMARY #length 917 #molecular_weight 103106 #checksum 8952
Query Match Best Local Similarity 69.3%; Score 52; DB 1; Length 917; Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 729 AYLTSSLDY 739
:|:|:|:|:|:|
QY 2 YLSTSSLDY 12
RESULT 5
ENTRY C71439 #type complete
TITLE hypothetical protein - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear cress
DATE 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
ACCESSIONS C71439
REFERENCE A71400
#authors Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirkse, W.; Van Staveren, M.; Stiekema, W.; Drost, L.; Ridley, P.; Hudson, S.A.; Patel, K.; Murphy, G.; Piffanelli, P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terryn, N.; Gielem, J.; Villarroel, R.; De Clerck, R.; Van Montagu, M.; Lecharny, A.; Auborg, S.; Gy, I.; Kreis, M.; Lao, N.; Kavanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger,

M.: Schaeffer, M.; Funk, B.; Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puidomenech, P.; Douka, A.; Vouklatou, E.; Milioni, D.; Hatzopoulos, P.; Piravandi, E.; Obermayer, B.; Hilbert, H.; Duesterhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palmer, K.; Benes, V.; Reckman, S.; Ansoorge, W.; Cooke, R.; Berger, C.; Delsen, M.; Voelt, M.; Volckaert, G.; Mewes, H.W.; Klosterman, S.; Schueller, C.; Chaiwatzis, N.
Nature (1998) 391:485-488
Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana.

#journal
#title
#cross-references MUID:98121113
#accession C71439
#status preliminary; nucleic acid sequence not shown;
translation not shown

##molecule_type DNA
##residues 1-471 ##label BEV
##cross-references GB:Z97342; NID:q2245031; PID:e327038; PID:g2245065

GENETICS
#map_position 4COP9-4G3845
#length 471 #molecular-weight 52785 #checksum 7455

Query Match
Best Local Similarity 68.0%; Score 51; DB 2; Length 471;
Pred. No. 1.05e+01;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 19 AGYLTSSDLD 29
1:1111111111
1 ASYLTSSDLD 11

RESULT 6
ENTRY T00040 #type complete
TITLE BH-protocadherin PCDH7 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change

ACCESSIONS
REFERENCE T00040
214074
Yoshida, K.; Yoshitomo-Nakagawa, K.; Seki, N.; Sasaki, M.; Sugano, S.
Genomics (1998) 49:458-461
Cloning, expression analysis, and chromosomal localization of BH-protocadherin (PCDH7), a novel member of the cadherin superfamily.

GENETICS
#accession T00040
#status preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-1069 ##label YOS
##cross-references EMBL:AB006755; NID:d1184677; PID:d1026122
##experimental_source clone BH-Pcdh-a

SUMMARY
#map_position 4P15
#length 1069 #molecular-weight 116104 #checksum 9974

Query Match
Best Local Similarity 68.0%; Score 51; DB 2; Length 1069;
Pred. No. 1.05e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 486 FLHTSTPLDY 495
1:1111111111
3 YLTSTSSLDY 12

RESULT 7
ENTRY T00043 #type complete
TITLE BH-protocadherin-a - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change

ACCESSIONS
REFERENCE T00043
214075
Yoshida, K.

#submission submitted to the EMBL Data Library, August 1997
#accession T00043
#status preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-1069 ##label YOS
##cross-references EMBL:AB006758; NID:d1227200; PID:d1033562

GENETICS
#gene pcdh7
#map_position 5C3-D
#length 1069 #molecular-weight 116313 #checksum 4821

SUMMARY
#length 1069 #molecular-weight 116313 #checksum 4821

Query Match
Best Local Similarity 68.0%; Score 51; DB 2; Length 1069;
Pred. No. 1.05e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 486 FLHTSAPLDY 495
1:1111111111
3 YLTSTSSLDY 12

RESULT 8
ENTRY T00041 #type complete
TITLE BH-protocadherin PCDH7 (clone BH-Pcdh-b) - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change

ACCESSIONS
REFERENCE T00041
214074
Yoshida, K.; Yoshitomo-Nakagawa, K.; Seki, N.; Sasaki, M.; Sugano, S.
Genomics (1998) 49:458-461
Cloning, expression analysis, and chromosomal localization of BH-protocadherin (PCDH7), a novel member of the cadherin superfamily.

GENETICS
#accession T00041
#status preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-1072 ##label YOS
##cross-references EMBL:AB006756; NID:d1184678; PID:d1026123
##experimental_source clone BH-Pcdh-b

SUMMARY
#map_position 4P15
#length 1072 #molecular-weight 116462 #checksum 9727

Query Match
Best Local Similarity 68.0%; Score 51; DB 2; Length 1072;
Pred. No. 1.05e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 486 FLHTSTPLDY 495
1:1111111111
3 YLTSTSSLDY 12

RESULT 9
ENTRY T00042 #type complete
TITLE BH-protocadherin PCDH7 (clone BH-Pcdh-c) - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change

ACCESSIONS
REFERENCE T00042
214074
Yoshida, K.; Yoshitomo-Nakagawa, K.; Seki, N.; Sasaki, M.; Sugano, S.
Genomics (1998) 49:458-461
Cloning, expression analysis, and chromosomal localization of BH-protocadherin (PCDH7), a novel member of the cadherin superfamily.

GENETICS
#accession T00042
#status preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-1200 ##label YOS
##cross-references EMBL:AB006757; NID:d1184679; PID:d1026124
##experimental_source clone BH-Pcdh-c

GENETICS

map-position 4p15
length 1200 #molecular-weight 130337 #checksum 7152

Query Match
Best Local Similarity 60.0%; Score 51; DB 2; Length 1200;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 439 FLHSTPDL 448
OY 3 YLSTSSSLDY 12

RESULT 10
ENTRY D70108 #type complete
TITLE conserved hypothetical protein BB0068 - Lyme disease
ORGANISM Spirochaete
#formal_name Borrelia burgdorferi #common_name Lyme disease
#spirochaete
13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change

ACCESSIONS
REFERENCE D70108
#authors Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.;

Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.;
Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb,
J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.;
Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.;
Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman,
J.; Uitterlinden, T.; Matthey, L.; McDonald, L.; Artach, P.;
Bowman, C.; Garland, S.; Fujita, C.; Cotton, M.D.; Horst,
K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.

#journal Nature (1997) 390:580-586
#title Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.

#cross-references MUID:98065943

#accession D70108

#status preliminary; nucleic acid sequence not shown;
translation not shown

#molecule_type DNA

#residues 1-293 #label KLE

#cross-references GB:AE001120; GB:AE000783; NID:g2687951; PID:g2687956;

SUMMARY #experimental_source strain B31
#length 293 #molecular-weight 33278 #checksum 5223

Query Match
Best Local Similarity 60.0%; Score 50; DB 2; Length 293;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

DB 198 AYLSTPNSLE 207
OY 2 YLSTSSSLDY 11

RESULT 11
ENTRY D70048 #type complete
TITLE ABC transporter (amino acid permease) homolog yvsh - *Bacillus subtilis*
ORGANISM #formal_name *Bacillus subtilis*
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change

ACCESSIONS
REFERENCE D70048
#authors Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;

Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;
Biotin, A.; Borchert, S.; Boriss, R.; Bouslier, L.; Brans,
A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
Choi, S.K.; Codani, J.J.; Conerton, I.F.; Cummings, N.J.;
Daniel, R.A.; Denizot, F.; Devine, K.M.; Diesterheft, A.;
Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.;
Fabret, C.; Ferrari, E.; Foulger, D.; Filtz, C.; Fujita,

M.; Fujita, Y.; Funa, S.; Galizzi, A.; Galleron, N.; Ghim,
S.Y.; Glaser, P.; Goffeau, A.; Gollightly, E.J.; Grandi, G.;
Guix, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.;
Hulo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.;
Kashara, Y.; Kraerr-Blanchard, M.; Klein, C.; Kobayashi,
Y.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.;
Kunita, K.; Lapidus, A.; Lardinis, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
M.; Moestl, D.; Nakai, S.; Noback, B.; Noone, D.; O'Reilly,
V.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro,
V.; Pohl, T.M.; Portecaille, D.; Porwollik, S.; Prescott,
A.M.; Prescan, E.; Puig, P.; Purnelle, B.; Rapoport, G.;
Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, E.;
Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo,
B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;
Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.;
Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;
Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.;
Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenger, T.;
Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumbstein, E.;
Yoshikawa, H.; Danchin, A.

#journal Nature (1997) 390:249-256
#title The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

#cross-references MUID:98044033

#accession D70048

#status preliminary; nucleic acid sequence not shown;
translation not shown

#molecule_type DNA

#residues 1-469 #label KUN

#cross-references GB:299121; GB:AL009126; NID:g2635827; PID:el186022;

PID:g2635847

#experimental_source strain 168

GENETICS

#gene yvsh

CLASSIFICATION #superfamily L-lysine transport protein

SUMMARY #length 469 #molecular-weight 50258 #checksum 4200

Query Match
Best Local Similarity 36.4%; Score 50; DB 2; Length 469;
Matches 4; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

DB 376 TFLTRAILAY 386
OY 2 YLSTSSSLDY 12

RESULT 12
ENTRY IKECB #type complete
TITLE colicin Ib - *Escherichia coli* plasmid ColIb
ORGANISM #formal_name *Escherichia coli*
DATE 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change

ACCESSIONS
REFERENCE A93533; D25035; B22503; A03503
#authors Vartley, J.M.; Boulnois, G.J.
#journal Nucleic Acids Res. (1984) 12:6727-6739
#title Analysis of a cloned colicin IB gene: complete nucleotide
sequence and implications for regulation of expression.

#cross-references MUID:85014128

#accession A93533

#molecule_type DNA
#residues 1-626 #label VAR
#cross-references GB:X01009; NID:941141; PID:941142

#note the authors translated the codon GAA for residue 294 as
Gln

REFERENCE A91822
#authors Mankovich, J.A.; Hsu, C.H.; Konisky, J.

#journal J. Bacteriol. (1986) 168:228-236
#title DNA and amino acid sequence analysis of structural and
#cross-references MUID:87008385
#accession D25035
##molecule-type DNA
##residues 1-626 ##label MAN
REFERENCE A22503
#authors Mankovich, J.A.; Lai, P.H.; Gokul, N.; Konisky, J.
#journal J. Biol. Chem. (1984) 259:8764-8768
#title Organization of the colicin IB gene.
#cross-references MUID:84264487
#accession A22503
##molecule-type DNA
##residues 1-40 ##label MA2
#accession B22503
##molecule-type protein
##residues 2-21 ##label MA3
REFERENCE A93546
#authors Varley, J.M.; Boulinois, G.J.
#journal Nucleic Acids Res. (1984) 12:8748
#contents annotation: corrigendum
COMMENT This bacteriocidal protein functions by depolarizing the cytoplasmic
membrane of sensitive cells.
GENETICS
#gene cib
#genome plasmid
CLASSIFICATION #superfamily colicin IB
KEYWORDS antibiotic; bacteriocin; toxin; transmembrane protein
FEATURE
SUMMARY #length 626 #molecular-weight 69923 #checksum 8113
Query Match 66.7%; Score 50; DB 1; Length 626;
Best Local Similarity 63.6%; Pred. No. 1.64e+01;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 333 NYLTHSGLDY 343
:|:|:|:|:|
QY 2 SYLSTSSLDY 12
RESULT 13
ENTRY I56333 #type fragment
TITLE apolipoprotein B - rat (fragment)
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
23-Feb-1997
SESSIONS I56333
REFERENCE I56333
#authors Reuben, M.A.; Svenson, K.L.; Doolittle, M.H.; Johnson, D.F.;
Lusis, A.J.; Elvovson, J.
#journal J. Lipid Res. (1988) 29:1337-1347
#title Biosynthetic relationships between three rat apolipoprotein B
peptides.
#cross-references MUID:89176719
#accession I56333
##status preliminary; translated from GB/EMBL/DBJ
##molecule-type mRNA
##residues 1-989 ##label RES
##cross-references GB:M27440; NID:g623548; PID:g623549
GENETICS
#gene apob
CLASSIFICATION #superfamily apolipoprotein B
SUMMARY #length 989 #checksum 1918
Query Match 66.7%; Score 50; DB 2; Length 989;
Best Local Similarity 60.0%; Pred. No. 1.64e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 175 YLQASTSLHY 184
:|:|:|:|
QY 3 YLSTSSLDY 12

RESULT 14
ENTRY T03180 #type complete
TITLE tyrosine protein kinase homolog - Chilo iridescent virus
ORGANISM #formal_name Chilo iridescent virus
DATE 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change
24-Mar-1999
ACCESSIONS T03180
REFERENCE 214834
#authors Bahr, U.; Tidona, C.A.; Darai, G.
#journal Virus Genes (1997) 15:235-245
#title The DNA sequence of Chilo iridescent virus between the genome
coordinates 0.101 and 0.391: similarities in coding
strategy between insect and vertebrate iridoviruses.
#accession T03180
##status preliminary; translated from GB/EMBL/DBJ
##molecule-type DNA
##residues 1-1186 ##label BAH
SUMMARY #cross-references EMBL:AF003534; NID:g2738385; PID:g2738451
#length 1186 #molecular-weight 138020 #checksum 6252
Query Match 66.7%; Score 50; DB 2; Length 1186;
Best Local Similarity 63.6%; Pred. No. 1.64e+01;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Db 503 TYLSTSSLDY 513
:|:|:|:|:|
QY 2 SYLSTSSLDY 12
RESULT 15
ENTRY A24571 #type complete
TITLE protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor -
human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 25-Oct-1987 #sequence_revision 06-Dec-1996 #text_change
22-May-1998
ACCESSIONS A24571; A25491; A44188; B44188; I59509; I57622
REFERENCE A24571
#authors Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.;
Miyajima, N.; Saito, T.; Toyoshima, K.
#journal Nature (1986) 319:230-234
#title Similarity of protein encoded by the human c-erb-B-2 gene to
epidermal growth factor receptor.
#cross-references MUID:86118663
#accession A24571
##molecule-type mRNA
##residues 1-1255 ##label YAM
##cross-references GB:X03363; NID:g31197; PID:g31198
REFERENCE A25491
#authors Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
#journal Proc. Natl. Acad. Sci. U.S.A. (1985) 82:6497-6501
#title A v-erbB-related protooncogene, c-erbB-2, is distinct from
the c-erbB-1/epidermal growth factor receptor gene and is
amplified in a human salivary adenocarcinoma.
#cross-references MUID:86016729
#accession A25491
##molecule-type DNA
##residues 737-1031 ##label SEM
REFERENCE A44188
#authors Consens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray,
A.; McGrath, J.; Seeburg, P.H.; Libermann, T.A.;
Schlessinger, J.; Francke, U.; Levinson, A.; Ullrich, A.
#journal Science (1985) 230:1132-1139
#title Tyrosine kinase receptor with extensive homology to EGF
receptor shares chromosomal location with neu oncogene.
#cross-references MUID:86070181
#accession A44188
##molecule-type DNA

```

##residues 740-910 ##label COU1
##cross-references GB:M12036; NID:g183988; PID:g183989
#accession B44188
##molecule_type mRNA
##residues 1-517,'RALL','522','S','524-654','V','656-1169','A','1171-1255
##label COU2
##cross-references GB:M1730; NID:g183986
159509
REFERENCE
#authors King, C.R.; Kraus, M.H.; Aaronson, S.A.
#journal Science (1985) 229:974-976
#title Amplification of a novel v-erbB-related gene in a human
#cross-references M01D:85272597
#accession 159509
#status translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 832-909 ##label REX
##cross-references GB:L29395; NID:g459807; PID:g459808
157622
REFERENCE
#authors Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger,
J.; Givol, D.
#journal Mol. Cell. Biol. (1987) 7:2597-2601
#title Human HER2 (neu) promoter: evidence for multiple mechanisms
#cross-references M01D:87266898
#accession 157622
#status translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 1-191 ##label TAL
##cross-references GB:M16792; NID:g183983; PID:g553332
COMMENT Amplification and overexpression of this erbB-related gene occurs
in about 30% of human breast and ovarian cancers.

GENETICS
#gene GDB:ERBB2; NGL; NEU; HER-2
##cross-references GDB:120613; OMIM:164870
#map_position 17q21.1-17q21.1
#introns 25/1; 75/3; 147/1; 883/3
#note the 1st of introns is incomplete
FUNCTION
#description catalyzes the phosphorylation of a peptidyl tyrosine residue
by ATP
CLASSIFICATION
#superfamily epidermal growth factor receptor: protein kinase
homology
KEYWORDS
ATP; autophosphorylation; duplication; glycoprotein;
phosphoprotein; phosphotransferase; proto-oncogene;
receptor; transforming protein; transmembrane protein;
tyrosine-specific protein kinase

FEATURE
1-21
22-1255
#domain signal sequence #status predicted #label SIG\
#product protein-tyrosine kinase erbB2 #status predicted
#label MAT\
22-653 #domain extracellular #status predicted #label EXT\
70-304 #domain EGF receptor extracellular domain repeat #label
E1\
395-605 #domain EGF receptor extracellular domain repeat #label
EE2\
654-675 #domain transmembrane #status predicted #label TM\
676-1255 #domain intracellular #status predicted #label INT\
718-983 #domain protein kinase homology #label KIN\
726-734 #region protein kinase ATP-binding motif\
68,124,187,259,530,
571,629 #binding_site carbohydrate (Asn) (covalent) #status
predicted\
686 #binding_site phosphate (Thr) (covalent) (by protein
kinase C) #status predicted\
753 #active_site Lys #status predicted\
1248 #binding_site phosphate (Tyr) (covalent) (by
autophosphorylation) #status predicted
SUMMARY
#length 1255 #molecular-weight 137909 #checksum 9382

Query Match 66.7%; Score 50; DB 1; Length 1255;
Best Local Similarity 45.5%; Pred. NO. 1.64e+01;

```

```

Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Db 63 TYLPTNASTSF 73
:|:|:|:|:
QY 2 SYLSTSSSLDY 12

```

Search completed: Thu Sep 2 12:31:23 1999
Job time : 19 secs.

OY 1 ASYLSTSSLDY 12

RESULT 2
ID DHSD_HUMAN STANDARD: PRT: 159 AA.
AC 014521.
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE SUCCINATE DEHYDROGENASE [UBIQUINONE] CYTOCHROME B SMALL SUBUNIT
DE SUBUNIT (CIBS) (SUCCINATE-UBIQUINONE REDUCTASE MEMBRANE ANCHOR
DE SUBUNIT).
GN SDHD OR SDH4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
OC PRIMATES: CATARRHINI: HOMINIDAE: HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RT TISSUE-LIVER:
RT MEDLINE: 98194224.
RT HIRAWAKE H., TANIMAKI M., KIJIMA S., KITTA K.;
RT "Cytochrome b in human complex II (succinate-ubiquinone
RT oxidoreductase): cDNA cloning of the components in liver mitochondria
RT and chromosome assignment of the genes for the large (SDHC) and small
RT (SDHD) subunits to 1q21 and 11q23.";
RT CYTOGENET. CELL. GENET. 79:132-138(1997).
CC -1- SUBUNIT: COMPOSED OF A 27 KD IRON PROTEIN (IP), A 70 KD
CC FLAVOPROTEIN (FP) AND A CYTOCHROME B COMPOSED OF TWO INTEGRAL
CC MEMBRANE PROTEINS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB006202; D1022913; -.
DR MIM: 602690; -.
KW TRICARBOXYLIC ACID CYCLE; ELECTRON TRANSPORT; HEME; TRANSMEMBRANE;
KW MITOCHONDRION; TRANSIT PEPTIDE.
FT TRANSIT 1 56 MITOCHONDRION (POTENTIAL).
FT CHAIN 57 159 SUCCINATE DEHYDROGENASE [UBIQUINONE]
FT CYTOCHROME B SMALL SUBUNIT.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 126 142 POTENTIAL.
SQ SEQUENCE 159 AA; 17043 MW; F4221825 CRC32;
Query Match 72.0%; Score 54; DB 1; Length 159;
Best Local Similarity 50.0%; Pred. No. 9.05e-01;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

EX MEDLINE: 92070466.
RA HUTTON M.L., HARVEY R.J., BARNARD E.A., DARLISON M.G.;
RA "Cloning of a cDNA that encodes an invertebrate glutamate receptor
RA subunit.";
RT FEBS LETT. 292:111-114(1991).
CC -1- FUNCTION: L-GLUTAMATE ACTS AS AN EXCITATORY NEUROTRANSMITTER AT
CC MANY SYNAPSES IN THE CENTRAL NERVOUS SYSTEM. THE POSTSYNAPTIC
CC ACTIONS OF GLU ARE MEDIATED BY A VARIETY OF RECEPTORS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X60086; G9629; -.
DR PIR: S15681; ACGAE.
DR PIR: S18443; S18443.
DR PFAM: PF00060; 119 Chan: 1.
KW RECEPTOR, POSTSYNAPTIC MEMBRANE; IONIC CHANNEL; GLYCOPROTEIN; SIGNAL;
KW TRANSMEMBRANE.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 917 GLUTAMATE RECEPTOR.
FT DOMAIN 21 558 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 559 578 POTENTIAL.
FT TRANSMEM 599 617 POTENTIAL.
FT TRANSMEM 628 617 POTENTIAL.
FT TRANSMEM 819 839 POTENTIAL.
FT CARBOHYD 62 62 POTENTIAL.
FT CARBOHYD 95 95 POTENTIAL.
FT CARBOHYD 121 121 POTENTIAL.
FT CARBOHYD 125 125 POTENTIAL.
FT CARBOHYD 229 229 POTENTIAL.
FT CARBOHYD 251 251 POTENTIAL.
FT CARBOHYD 261 261 POTENTIAL.
FT CARBOHYD 272 272 POTENTIAL.
FT CARBOHYD 418 418 POTENTIAL.
FT CARBOHYD 419 419 POTENTIAL.
FT CARBOHYD 424 424 POTENTIAL.
FT CARBOHYD 491 491 POTENTIAL.
FT CARBOHYD 775 775 POTENTIAL.
FT CARBOHYD 881 881 POTENTIAL.
SQ SEQUENCE 917 AA; 103139 MW; 879CBEDC CRC32;
Query Match 69.3%; Score 52; DB 1; Length 917;
Best Local Similarity 54.5%; Pred. No. 2.48e+00;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 729 AYLSTSSLDY 739
OY 2 SYLSTSSLDY 12

RESULT 4
ID SOG_DROME STANDARD: PRT: 1038 AA.
AC 024025;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE DORSAL-VENTRAL PATTERNING PROTEIN SOG (SHORT GASTRULATION PROTEIN).
GN DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA: METAZOA: ARTHROPODA: TRACHEATA: HEXAPODA; INSECTA;
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYROIDEA;
OC DROSOPHILIDAE; DROSOPHILA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95047398.
RX FRANCOIS V., SOLLOWAY M., O'NEILL J.W., EMERY J., BIER E.;

RT "Dorsal-ventral patterning of the Drosophila embryo depends on a putative negative growth factor encoded by the short gastrulation gene".

RT GENES DEV. 8:2602-2616(1994).

CC -1- FUNCTION: PUTATIVE NEGATIVE GROWTH FACTOR; ANTAGONIST OF DPP, A PROTEIN INVOLVED IN PATTERNING THE DORSAL REGION AND IN THE DEVELOPMENT OF THE NEUROECTODERM.

CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).

CC -1- TISSUE SPECIFICITY: ABUTS THE DORSAL DPP-EXpressING CELLS IN A LATERAL STRIPE 14-16 CELLS WIDE. LATER IN EMBRYOGENESIS IT IS EXPRESSED IN NEUROECTODERM AND IN THE ENDODERM SPACED ALONG THE ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING GUT.

CC -1- DEVELOPMENTAL STAGE: EMBRYOGENESIS.

CC -1- SIMILARITY: TO XENOPUS DORSALIZING FACTOR CHORDIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U18774; G1203794; -.

CC DR FLYBASE: FBgn0003463; sog.

CC DR PROSITE: PS01208; VWFc; 2.

CC DR PRAM: PF00093; VWFc; 4.

CC KW TRANSMEMBRANE; DEVELOPMENTAL PROTEIN; REPEAT; GROWTH FACTOR; GROWTH REGULATION; SIGNAL-ANCHOR.

CC FT DOMAIN 1 53 CYTOPLASMIC (POTENTIAL).

CC FT TRANSMEM 54 74 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)

CC FT FT (POTENTIAL).

CC FT DOMAIN 75 1038 EXTRACELLULAR (POTENTIAL).

CC FT REPEAT 100 175 VWFc 1.

CC FT REPEAT 421 522 SR1.

CC FT REPEAT 592 668 SR2.

CC FT REPEAT 677 754 SR3.

CC FT DOMAIN 742 804 VWFc 2.

CC FT DOMAIN 830 899 VWFc 3.

CC FT DOMAIN 939 1020 VWFc 4.

CC FT CARBOHYD 179 179 POTENTIAL.

CC FT CARBOHYD 287 287 POTENTIAL.

CC FT CARBOHYD 520 520 POTENTIAL.

CC FT CARBOHYD 666 666 POTENTIAL.

CC FT CARBOHYD 752 752 POTENTIAL.

CC FT CARBOHYD 821 821 POTENTIAL.

CC FT SEQUENCE 1038 AA; 115514 MW; DCADAEFF5 CRC32;

Query Match 68.0%; Score 51; DB 1; Length 1038;

Best Local Similarity 60.0%; Pred. No. 4.06e+00;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 497 YLNTDGLAY 506

QY 3 YLSTSSLDY 12

RESULT 5 STANDARD; PRT; 626 AA.

ID CEIB_ECOLI

AC P04479;

DT 13-AUG-1987 (REL. 05, CREATED)

DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE COLICIN IB PROTEIN.

GN CIB.

OS ESCHERICHIA COLI.

OC PLASMID INCII COLIB-P9

OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;

OC ESCHERICHIA.

OC [1]

RN SEQUENCE FROM N.A.

RX MEDLINE; 87008385.

RA MANKOVICH J.A., HSU C.-H., KONISKY J.;

RT "DNA and amino acid sequence analysis of structural and immunity genes of colicins Ia and Ib".

RT J. BACTERIOL. 168:228-236(1986).

CC [2]

CC SEQUENCE OF 1-40 FROM N.A.

CC RX MEDLINE; 84264487.

CC RA MANKOVICH J.A., LAI P.H., GOKUL N., KONISKY J.;

CC RT "Organization of the colicin Ib gene. Promoter structure and immunity domain".

CC RT J. BIOL. CHEM. 259:8764-8768(1984).

CC [3]

CC SEQUENCE FROM N.A.

CC RX MEDLINE; 85014128.

CC RA VARLEY J.M., BOUNOIS G.J.;

CC RT "Analysis of a cloned colicin Ib gene: complete nucleotide sequence and implications for regulation of expression".

CC RT NUCLEIC ACIDS RES. 12:6727-6739(1984).

CC [4]

CC ERRATUM.

CC RA VARLEY J.M., BOUNOIS G.J.;

CC RT NUCLEIC ACIDS RES. 12:8748-8748(1984).

CC CC -1- FUNCTION: THIS COLICIN IS A CHANNEL-FORMING COLICIN. THIS CLASS OF TRANSMEMBRANE TOXINS DEPOLARIZE THE CYTOPLASMIC MEMBRANE, LEADING TO DISSIPATION OF CELLULAR ENERGY.

CC CC -1- FUNCTION: COLICINS ARE POLYPEPTIDE TOXINS PRODUCED BY AND ACTIVE AGAINST, ESCHERICHIA COLI AND CLOSELY RELATED BACTERIA.

CC CC -1- SIMILARITY: BELONGS TO THE CHANNEL FORMING COLICIN FAMILY.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC CC -----

CC EMBL; K02071; G144640; -.

CC DR EMBL; X01009; G41142; -.

CC DR EMBL; M13820; G144648; -.

CC DR PIR; A03503; IRECB.

CC DR PIR; A22503; A22503.

CC DR PIR; D25035; D25035.

CC DR PROSITE: PS00276; CHANNEL-COLICIN; 1.

CC DR PRAM: PF01024; Colicin; 1.

CC DR HSSP; P06716; ICII.

CC KW PLASMID; BACTERIOCIN; COLICIN; TOXIN; TRANSMEMBRANE.

CC FT TRANSMEM 588 612 POTENTIAL.

CC FT SEQUENCE 626 AA; 69923 MW; 983D11B0 CRC32;

Query Match 66.7%; Score 50; DB 1; Length 626;

Best Local Similarity 63.6%; Pred. No. 6.61e+00;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 333 NYLTHSGLDY 343

QY 2 YLSTSSLDY 12

RESULT 6 STANDARD; PRT; 1255 AA.

ID ERB2_HUMAN

AC P04626;

DT 13-AUG-1987 (REL. 05, CREATED)

DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE ERBB-2 RECEPTOR PROTEIN-TYROSINE KINASE PRECURSOR (EC 2.7.1.112)

DE (P185ERBB2) (NEU PROTO-ONCOGENE) (C-ERBB-2).

GN ERBB2 OR HER2 OR NCL OR NEU.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.

OC [1]

RN SEQUENCE FROM N.A.

RX MEDLINE: 86118663.
 RA YAMAMOTO T., IKAWA S., AKIYAMA T., SEMBA K., NOMURA N., MIYAJIMA N.,
 RA SAITO T., TOYOSHIMA K.,
 RT "Similarity of protein encoded by the human c-erb-B-2 gene to
 RT epidermal growth factor receptor.",
 RL NATURE 319:230-234(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 86070181.
 RA COUSSENS L., YANG-FENG T.L., LIAO Y.C., CHEN E., GRAY A.,
 RA MCGRATH J., SEEBURG P.H., LIBERMANN T.A., SCHLESSINGER J.,
 RA FRACKE U., LEVINSON A., ULLRICH A.,
 RT "Tyrosine kinase receptor with extensive homology to EGF receptor
 RT shares chromosomal location with neu oncogene.",
 RL SCIENCE 230:1132-1139(1985).
 RN [3]
 RP SEQUENCE OF 737-1031 FROM N.A.
 RX MEDLINE: 86016729.
 RA SEMBA K., KAWATA N., TOYOSHIMA K., YAMAMOTO T.,
 RA "A v-erbB-related protooncogene, c-erbB-2, is distinct from the
 RA c-erbB-1/epidermal growth factor-receptor gene and is amplified in a
 RA human salivary gland adenocarcinoma.",
 RL PROC. NATL. ACAD. SCI. U.S.A. 82:6497-6501(1985).
 CC -1- FUNCTION: NEUREGULINS AND GP30 ARE POTENTIAL LIGANDS FOR THIS
 CC RECEPTOR. NOT ACTIVATED BY EGF, TGF-ALPHA AND AMPHIREGULIN.
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
 CC PROTEIN TYROSINE PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
 CC (POTENTIAL).
 CC -1- PPM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 DR EMBL: M11767; G553282; -
 DR EMBL: M11761; G553282; JOINED.
 DR EMBL: M11762; G553282; JOINED.
 DR EMBL: M11763; G553282; JOINED.
 DR EMBL: M11764; G553282; JOINED.
 DR EMBL: M11765; G553282; JOINED.
 DR EMBL: M11766; G553282; JOINED.
 DR EMBL: M11730; G306840; -
 DR EMBL: M12036; G183989; -
 DR EMBL: X03363; G31198; -
 DR PIR: A25491; A25491.
 DR PIR: A24571; A24571.
 DR MIM: 164870; -
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR PFAM: PF00069; PKINASE; 1.
 DR PFAM: PF00757; Furin-like; 1.
 DR PFAM: PF01030; Recep_L-Domain; 2.
 DR HSSP: P11362; 1FG1.
 KW TRANSMEMBRANE: GLYCOPROTEIN; MULTIGENE FAMILY; RECEPTOR; SIGNAL;
 KW TRANSFERASE: TYROSINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION.
 FT SIGNAL 1 21
 FT CHAIN 22 1255
 FT DOMAIN 22 652
 FT TRANSMEM 653 675
 FT DOMAIN 676 1255
 FT DOMAIN 720 987
 FT NP_BIND 726 734
 FT BINDING 753 753
 FT ACT_SITE 845 845
 FT BY SIMILARITY.

FT MOD_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1248 1248 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 68 68 POTENTIAL.
 FT CARBOHYD 124 124 POTENTIAL.
 FT CARBOHYD 187 187 POTENTIAL.
 FT CARBOHYD 259 259 POTENTIAL.
 FT CARBOHYD 330 330 POTENTIAL.
 FT CARBOHYD 571 571 POTENTIAL.
 FT CARBOHYD 629 629 POTENTIAL.
 FT CARBOHYD 655 655 POTENTIAL.
 FT CONFLICT 1170 1170 I -> V (IN REF. 2).
 FT CONFLICT 1170 1170 P -> A (IN REF. 2).
 SQ SEQUENCE 1255 AA; 137909 MW; 715C377C CRC32;
 Query Match 66.7%; Score 50; DB 1; Length 1255;
 Best Local Similarity 45.5%; Pred. No. 6,61e+00;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 Db 63 TYLPNTASLSF 73
 Qy 2 SYLSTSSSLDY 12
 RESULT 7
 ID 5HT1_DROME STANDARD. PRT: 564 AA.
 AC P20905;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE 5-HYDROXYTRYPTAMINE RECEPTOR 1 (5-HT RECEPTOR) (SEROTONIN RECEPTOR).
 GN 5HT1-RL OR 5-HT7.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
 OC DROSOPHILIDAE; DROSOPHILA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON-R; TISSUE=HEAD;
 RX MEDLINE: 91062395.
 RA WITZ P., AMLAIKY N., PLASSAT J.-L., MAROTEAUX L., BORRELLI E., HEN R.,
 RT "Cloning and characterization of a Drosophila serotonin receptor that
 RT activates adenylylate cyclase.",
 RL PROC. NATL. ACAD. SCI. U.S.A. 87:8940-8944(1990).
 CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR
 CC 5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTION
 CC AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLYLATE
 CC CYCLASE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: HEAD.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC BUT WITH ONE EXTRA POTENTIAL TRANSMEMBRANE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 DR EMBL: M55533; G156725; -
 DR PIR: A38271; A38271.
 DR GCRDB: GCR_0023; -
 DR FLYBASE: FBgn0004573; 5-HT7.
 DR PROSITE: PS00337; G-PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT.
 FT TRANSMEM 29 51
 FT TRANSMEM 165 188
 FT TRANSMEM 189 198
 FT DOMAIN 199 222
 FT TRANSMEM 223 236
 FT DOMAIN 237 258
 FT TRANSMEM 258 258
 FT 3 (POTENTIAL).

```

FT DOMAIN 259 278 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 279 302 4 (POTENTIAL).
FT DOMAIN 303 330 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 331 353 5 (POTENTIAL).
FT DOMAIN 354 454 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 455 476 6 (POTENTIAL).
FT DOMAIN 477 487 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 488 510 7 (POTENTIAL).
FT DOMAIN 511 564 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 89 106 9 x 2 AA TANDEM REPEATS OF G-S.
FT DISULFID 235 314 BY SIMILARITY.
SQ SEQUENCE 564 AA: 60861 MW: 312369E8 CRC32:

Query Match 65.3%: Score 49; DB 1; Length 564;
Best Local Similarity 40.0%: Pred. No. 1.07e+01;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 261 YLAITKPLEY 270
||:|:|:|
3 YLSTSSLDY 12

RESULT 8 STANDARD: PRT: 1469 AA.
ID DP27_CAEEL
AC P48996;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CHROMOSOME CONDENSATION PROTEIN DPY-27.
DN DPY-27 OR R13G10.1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 95042743.
RA CHUANG P.-T., ALBERTSON D.G., MEYER B.J.;
RT "DPY-27: a chromosome condensation protein homolog that regulates C.
RT elegans dosage compensation through association with the X
RT chromosome."
RL CELL 79:459-474(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA GARDNER A.;
RL SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: ACTS TO REDUCE EXPRESSION OF BOTH HERMAPHRODITE X
CC CHROMOSOMES. DPY-27 BECOMES SPECIFICALLY LOCALIZED TO THE X
CC CHROMOSOMES OF WILD-TYPE XX EMBRYOS, BUT REMAINS DIFFUSELY
CC DISTRIBUTED THROUGHOUT THE NUCLEI OF MALE (XO) EMBRYOS. COULD
CC IMPLEMENTS DOSAGE COMPENSATION BY CONDENSING THE CHROMATIN
CC STRUCTURE OF X IN A MANNER THAT CAUSES GENERAL REDUCTION OF X
CC CHROMOSOME EXPRESSION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN EMBRYOS AND EARLY-STAGED LARVAE.
CC -1- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS
CC FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L35274; G529385; -
DR EMBL: Z35602; E1348662; -
DR MORPEP: R13610.1; CE01052.
KW ATP-BINDING; COILED COIL; NUCLEAR PROTEIN.
NP_BIND 122 129 ATP (POTENTIAL).

```

```

FT DOMAIN 875 916 COILED COIL (POTENTIAL).
FT TRANSMEM 1014 1171 COILED COIL (POTENTIAL).
FT DOMAIN 1253 1280 ALA/ASP-RICH (DA-BOX).
FT MOTAGEN 128 128 K->E: LOSS OF FUNCTION.
FT MOTAGEN 128 128 K->I: LOSS OF FUNCTION.
SQ SEQUENCE 1469 AA: 169618 MW: D08782B5 CRC32:

Query Match 65.3%: Score 49; DB 1; Length 1469;
Best Local Similarity 45.5%: Pred. No. 1.07e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 316 AYLTFFNNLY 326
||:|:|:|
2 YLSTSSLDY 12

RESULT 9 STANDARD: PRT: 37 AA.
ID YRVL_CAEEL
AC Q19177;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 4.1 KD PROTEIN F07H5.4 IN CHROMOSOME II.
DN F07H5.4.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA STEWARD C.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP REVISIONS.
RC STRAIN-BRISTOL N2;
RA JONES S.J.M.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z68314; E1345182; -
DR MORPEP: F07H5.4; CE09231.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 37 AA: 4082 MW: C16B5292 CRC32:

Query Match 64.0%: Score 48; DB 1; Length 37;
Best Local Similarity 60.0%: Pred. No. 1.77e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 2 YLDPSSLDY 11
||:|:|:|
3 YLSTSSLDY 12

RESULT 10 STANDARD: PRT: 272 AA.
ID YLGL_HAEIN
AC P44771;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN H10597.
DN H10597.
OS HAEMOPHILUS INFLUENZAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURILLACEAE;
OC HAEMOPHILUS.
RN [1]
RP SEQUENCE FROM N.A.

```

```
RC STAIN-RD / KW20.  
RX MEDLINE: 95350630.  
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,  
RA KERLAAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,  
RA MCKENNEY K., SUTTON G., FITZHUGH W., FILDS C.A., GOCAYNE J.D.,  
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLADKE A., KELLEY J.M.,  
RA WEIDMAN J.E., PHILLIPS C.A., SPRIGGS T., HEDBLUM E., COTTON M.D.,  
RA UTERERACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,  
RA FINE L.D., FITCHMAN J.L., FUHRMANN J.L., GEOGGAGEN N.S.M.,  
RA GENHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,  
RA VENTER J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus  
influenzae Rd.";  
RL SCIENCE 269:496-512(1995).  
CC -I- SIMILARITY: BELONGS TO THE COE/YBHA/YIDA/YIGL (E.COLI) / YGSE/YXEH  
CC (B.SUBTILIS) FAMILY. STRONG, TO E.COLI YIGL.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: U32741; GI573586; -.  
DR TIGR: HI0597; -.  
DR PROSITE: PS01228; COF_1; 1.  
DR PROSITE: PS01229; COF_2; 1.  
DR PFAM: PFO0592; DUF3; 1.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 272 AA; 30523 MW; 7553B65C CRC32;  
  
Query Match          64.0%; Score 48; DB 1; Length 272;  
Best Local Similarity 58.3%; Pred. No. 1.71e+01;  
Matches      7; Conservative    1; Mismatches     4; Indels      0; Gaps      0;  
  
Db   261 ARYLTOFGLDY 272  
      | | | : |||  
Qy   1 ASYSTSSSLDY 12  
  
RESULT  11  
AC       O08361; STANDARD;             PRY; 297 AA.  
DT       01-NOV-1997 (REL. 35, CREATED)  
DT       01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT       01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
ST       PHOSPHORIBOSYLAMINOIMIDAZOLE-SUCCINOCARBOXAMIDE SYNTHASE (EC 6.3.2.6)  
        (SAICAR SYNTHETASE).  
GN       PURC OR MLCSB.16.  
GC       MYCOBACTERIUM LEPRAE.  
OS       BACTERIA: FIRMICUTES: ACTINOBACTERIA: ACTINOBACTERIADE:  
OC       ACINOMYCETALES; CONRYNBACTERINEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.  
NC       [1]  
RP       SEQUENCE FROM N.A.  
RF       BACCOCK K., CHURCHER C.M., PARKHILL J., BARRELL B.G.,  
RA       RAJANDREAM M.A.;  
RL       SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDAJ DATA BANKS.  
CC       -I- CATALYTIC ACTIVITY: ATP + 1-(5'-PHOSPHORIBOSYL)-4-CARBOXY-5-  
CC       AMINOIMIDAZOLE + L-ASPARATE ~> ADP + ORTHOPHOSPHATE + 1-(5-  
CC       PHOSPHORIBOSYL)-4-(N-SUCCINO-CARBOXAMIDE)-5-AMINOIMIDAZOLE.  
CC       -I- PATHWAY: SEVENTH STEP IN DE NOVO PURINE BIOSYNTHESIS.  
CC       -I- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC       -I- SIMILARITY: TO OTHER SAICAR SYNTHETASES.  
CC       -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).
```

```

CC -----
DR EMBL; 295151; E315054; -
DR PROSITE; PS01057; SAICAR_SYNTHETASE_1; 1.
DR PROSITE; PS01058; SAICAR_SYNTHETASE_2; 1.
DR PFAM; PF01259; SAICAR_synt; 1.
KW PURINE BIOSYNTHESIS; LIGASE.
SO SEQUENCE 297 AA; 33163 MW; 56F60A12 CRC32;

Query Match 64.0%; Score 48; DB 1; Length 297;
Best Local Similarity 54.5%; Pred. No. 1,71e+01;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 104 GYTGSGLLDY 114
QY 2 SYLSTSSSLDY 12
      11: 1: 111
      2 SYLSTSSSLDY 12

RESULT 12
ID PURF_MYCTU STANDARD: PRT; 297 AA.
AC Q39566; P77904;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PHOSPHORIBOSYLAMINOIMIDAZOLE-SUCCINOCARBOXAMIDE SYNTHASE (EC 6.3.2.6)
DE (SAICAR SYNTHETASE).
GN PURC OR MYC359.24.
OS MYCOBACTERIUM TUBERCULOSIS.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIADAE;
OC ACTINOMYCETALES; CORYNEBACTERIINEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RC MEDLINE; 96425868.
RA JACKSON M., BERTHER F.-X., OTAL I., RAUZIER J., MARTIN C.,
RA GIGOUEL B., GUILHOT C.;
RT "The mycobacterium tuberculosis purine biosynthetic pathway:
RT isolation and characterization of the purC and purL genes.";
RL MICROBIOLOGY 142:2439-2447(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MCELAN J., HARRIS D., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- CATALYTIC ACTIVITY: ATP + 1-(5-PHOSPHORIBOSYL)-4-CARBOXY-5-
CC AMINOIMIDAZOLE + L-ASPARTATE ~ ADP + ORTHOPHOSPHATE + 1-(5-
CC PHOSPHORIBOSYL)-4-(N-SUCCINO-CARBOXAMIDE)-5-AMINOIMIDAZOLE.
CC -1- PATHWAY: SEVENTH STEP IN DE NOVO PURINE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMOTRIMER (BT SIMILARITY).
CC -1- SIMILARITY: TO OTHER SAICAR SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/,
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 280226; E266570; -
DR EMBL; U34957; G1144522; -
DR PROSITE; PS01057; SAICAR_SYNTHETASE_1; 1.
DR PROSITE; PS01058; SAICAR_SYNTHETASE_2; 1.
DR PFAM; PF01259; SAICAR_synt; 1.
KW PURINE BIOSYNTHESIS; LIGASE.
CC CONFLICT 124 124 A -> R (IN REF. 1).
CC CONFLICT 166 166 A -> P (IN REF. 1).
SQ SEQUENCE 297 AA; 32930 MW; 23A860DC CRC32;

Query Match 64.0%; Score 48; DB 1; Length 297;
Best Local Similarity 54.5%; Pred. No. 1,71e+01;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 104 GYTGSGLLDY 114
QY 2 SYLSTSSSLDY 12
      11: 1: 111
      2 SYLSTSSSLDY 12

Query Match 64.0%; Score 48; DB 1; Length 297;
Best Local Similarity 54.5%; Pred. No. 1,71e+01;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 104 GYTGSGLLDY 114
QY 2 SYLSTSSSLDY 12
      11: 1: 111
      2 SYLSTSSSLDY 12

```


QY 2 SYLSTSSLDY 12

RESULT 13

ID BAR_MELGA STANDARD: PRT: 428 AA.

AC P43141:

DT 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DE 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE BETA-4C ADRENERGIC RECEPTOR.

GN ADRA4C.

OS MELEAGRIS GALLOPAVO (COMMON TURKEY).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;

OC NEOGNATHAE; GALLIFORMES; MELEAGRIDIDAE; MELEAGRIS.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 95014249.

RA CHEN X.-H., HARDEN T.K., NICHOLAS R.A.;

RA "Molecular cloning and characterization of a novel beta-adrenergic receptor."

CC J. BIOL. CHEM. 269:24810-24819(1994).

CC -1- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-INDUCED ACTIVATION OF ADENYLYLATE CYCLASE THROUGH THE ACTION OF G PROTEINS.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- TISSUE SPECIFICITY: BROAD TISSUE DISTRIBUTION.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U13977; G556604; -

DR EMBL; U13978; G555882; -

DR PROSITE; PS00237; G_PROTEIN_RECEPTOR; 1.

DR PFAM; PF00001; 7tm_1; 1.

DR HSP; P07700; IDEP.

KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;

KW MULTIGENE FAMILY; PHOSPHORYLATION; LIPOPROTEIN; PALMITATE.

FT DOMAIN 1 25

FT TRANSSEM 26 49

FT DOMAIN 50 58

FT TRANSSEM 59 77

FT DOMAIN 78 97

FT TRANSSEM 98 119

FT DOMAIN 120 141

FT TRANSSEM 142 164

FT DOMAIN 165 189

FT TRANSSEM 190 211

FT DOMAIN 212 261

FT TRANSSEM 262 283

FT DOMAIN 284 294

FT TRANSSEM 295 315

FT DOMAIN 316 428

FT CARBOHYD 8

FT CARBOHYD 13

FT DISULFID 96 175

FT LIPID 329

FT SEQUENCE 428 AA; 47398 MW; 8B794FOC CRC32;

Query Match 64.0%; Score 48; DB 1; Length 428;

Best Local Similarity 40.0%; Pred. No. 1.71e+01;

Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 122 YLAITAPLOY 131

QY 3 YLSTSSLDY 12

RESULT 14

ID SECY MYCTU STANDARD: PRT: 441 AA.

AC P49426:

DT 15-JUL-1998 (REL. 36, CREATED)

DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE PREPROTEIN TRANSLOCASE SECY SUBUNIT.

GN SECY OR RV0732 OR MYO41.06.

OS MYCOBACTERIUM TUBERCULOSIS, AND MYCOBACTERIUM BOVIS.

OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIUM.

OC ACTINOMYCETALES; CORNYNBACTERIINEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.

RN [1]

RP SEQUENCE FROM N.A.

RC SPECIES=M.TUBERCULOSIS; STRAIN=H37RV;

RX MEDLINE; 98295987.

RA COLE S.T., BROSCHE R., PARKHILL J., GARNIER T., CHURCHER C., HARRIS D.,

RA GORDON S.V., EIGMEIER K., GAS S., BARRY C.E., III, TEKAYA F.,

RA BADCOCK R., BASHAM D., BROWN D., CHILLINGWORTH T., CONNOR R.,

RA DAVIES R., DEVLIN K., FELTWEILL T., GENTLES S., HAMLIN N., HOLROYD S.,

RA HORNSBY T., JAGELS K., KROGH A., MCLEAN J., MOULE S., MURPHY L.,

RA OLIVER S., OSBORNE J., ODAIL M.A., RAJANDREAM M.A., ROGERS J.,

RA RUTTER S., SEEGER K., SKELTON S., SQUARES S., SQUARES R., SOLSTON J.E.,

RA TAYLOR K., WHITEHEAD S., BARRELL B.G.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."

RL NATURE 393:537-544(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC SPECIES=M.BOVIS; STRAIN=BCG;

RA KIM J.K., CHOE Y.K.;

RA SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT. INTERACTS WITH SECY AND SECY TO ALLOW THE TRANSLOCATION OF PROTEINS ACROSS THE PLASMA MEMBRANE, BY FORMING PART OF A CHANNEL (BY SIMILARITY).

CC -1- SUBUNIT: ONE OF SEVEN SECRETORY PROTEINS (SECY-F & SECY) THAT COMPRISE THE PROKARYOTIC PROTEIN TRANSLOCATION APPARATUS (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- SIMILARITY: BELONGS TO THE SECY/SECY-ALPHA FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AL021958; E1253270; -

DR EMBL; U77912; G181627; -

DR PROSITE; PS00755; SECY_1; 1.

DR PROSITE; PS00756; SECY_2; 1.

DR PFAM; PF00344; secy; 1.

KW PROTEIN TRANSFERT; TRANSMEMBRANE; TRANSLOCATION.

FT TRANSSEM 18 38

FT TRANSSEM 57 77

FT TRANSSEM 78 98

FT TRANSSEM 124 144

FT TRANSSEM 157 177

FT TRANSSEM 180 200

FT TRANSSEM 215 235

FT TRANSSEM 272 292

FT TRANSSEM 318 338

FT TRANSSEM 382 402

FT TRANSSEM 404 424

FT SEQUENCE 441 AA; 47611 MW; 26E1FAC6 CRC32;

Query Match 64.0%; Score 48; DB 1; Length 441;

Best Local Similarity 50.0%; Pred. No. 1.71e+01;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 309 GTYLDSPSNLVY 320

QY 1 ASYLTSSSLDY 12

RESULT 15
ID YM61 YEAST STANDARD; PRT; 475 AA.

AC 003652;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 55.3 KD PROTEIN IN RAR1-SCJ1 INTERGENIC REGION.
GN YMR211W OR YMR8261.05
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST)
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.

[1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA DEDMAN K., BROWN D., BOWMAN S., BARRELL B.G., RAJANDREAM M.A.,
RA WALSH S.V.;
RA SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; Z49809; G854463; -
DR PFAM; PF00091; tubulin; 1.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 475 AA; 55312 MW; 18FA4F03 CRC32;

Query Match 64.0%; Score 48; DB 1; Length 475;
Best Local Similarity 50.0%; Pred. No. 1,71e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 321 YLTATITLGY 330
QY 3 YLTSSSLDY 12

Search completed: Thu Sep 2 12:31:47 1999
Job time : 8 secs.

 Nucleotide
 (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPerch_pp protein - protein database search, using Smith-Waterman algorithm
 on: Thu Sep 2 12:32:06 1999; MasPar time 4.45 Seconds
 Molecular output not generated. 147.067 Million cell updates/sec

Title: >US-08-599-226-29
 Description: (1-12) from US08599226.pep
 Perfect Score: 75
 Sequence: 1 ASYSTSSSLDY 12

Scoring table: PAM 150
 Gap 15

Searched: 179066 segs, 54579741 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: sptrembl
 1:sp:archaea 2:sp:bacteria 3:sp:fungi 4:sp:human
 5:sp:invertebrate 6:sp:mammal 7:sp:mhc 8:sp:organelle
 9:sp:phage 10:sp:plant 11:sp:rodent 12:sp:unclassified
 13:sp:vertebrate 14:sp:virus

Statistics: Mean 24.003; Variance 31.514; scale 0.762

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	55	73.3	649	3	060167	PROTEIN COMPLEX ASSEMB
2	54	72.0	349	5	017959	MO1B2.5 PROTEIN.
3	54	72.0	503	2	026074	PROTEIN-EXPORT MEMBRAN
4	53	70.7	127	14	011696	NUCLEOPROTEIN (FRAGMEN
5	53	70.7	265	5	025144	HROX1
6	53	70.7	897	5	017336	LET 858.
7	53	70.7	1238	5	061198	F15E6.6 PROTEIN.
8	52	69.3	372	5	P91143	SIMILAR TO ACETYLTTRANS
9	52	69.3	475	13	093514	AXIAL PROTOCADHERIN (F
10	52	69.3	932	5	001623	SIMILAR TO LIGAND-GATE
11	51	68.0	83	9	048385	ORF83.
12	51	68.0	471	10	023552	HYPOTHETICAL 52.8 KD P
13	51	68.0	582	3	074931	ALTERNATIVE NADH-DEHYD
14	51	68.0	1035	13	057537	NF-PROTOCADHERIN-A.
15	51	68.0	1069	11	088185	BH-PROTOCADHERIN-A.
16	51	68.0	1069	4	060245	PCDH7 (BH-PCDH)A.
17	51	68.0	1072	4	060246	PCDH7 (BH-PCDH)B.
18	51	68.0	1200	4	060247	PCDH7 (BH-PCDH)C.
19	50	66.7	293	2	051095	CONSERVED HYPOTHETICAL
20	50	66.7	444	5	P91141	SIMILAR TO ACETYLTTRANS

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
21	50	66.7	469	2	032204	YVSH PROTEIN.
22	50	66.7	626	2	046732	COLICIN PROTEIN.
23	50	66.7	626	2	046734	COLICIN PROTEIN.
24	50	66.7	626	2	046738	COLICIN PROTEIN.
25	50	66.7	626	2	046740	COLICIN PROTEIN.
26	50	66.7	989	11	063052	APOLIPROTEIN B (FRAG
27	50	66.7	1186	14	055767	PUTATIVE TYROSINE PROT
28	49	65.3	112	3	007255	INTERNAL ORF OF L3149
29	49	65.3	204	3	012097	HYPOTHETICAL 22.6 KD P
30	49	65.3	222	13	P70023	OLEFACTORY RECEPTOR (FR
31	49	65.3	377	2	084397	HYPOTHETICAL 41.4 KD P
32	49	65.3	384	2	011197	HYPOTHETICAL 41.9 KD P
33	49	65.3	392	2	050873	CONSERVED HYPOTHETICAL
34	49	65.3	409	2	025853	NADH-UBIQUINONE OXIDOR
35	49	65.3	650	2	050417	MULTI-FUNCTIONAL ENZYM
36	49	65.3	1026	4	008174	PROTOCADHERIN 42 PRECU
37	49	65.3	1406	4	015082	KIAA0377.
38	48	64.0	170	2	086052	AMMONIUM TRANSPORTER (
39	48	64.0	325	2	055392	HYPOTHETICAL 35.4 KD P
40	48	64.0	420	3	P78765	F15SION YEAST (FRAGMEN
41	48	64.0	438	2	067997	AMTB.
42	48	64.0	505	11	088808	TUBBY PROTEIN.
43	48	64.0	954	2	084584	HYPOTHETICAL 109.2 KD
44	48	64.0	966	11	055098	SERINE/THREONINE KINAS
45	48	64.0	1685	10	004142	RNA POLYMERASE II LARG

ALIGNMENTS

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	55	73.3	649	3	060167	PROTEIN COMPLEX ASSEMB
2	54	72.0	349	5	017959	MO1B2.5 PROTEIN.
3	54	72.0	503	2	026074	PROTEIN-EXPORT MEMBRAN
4	53	70.7	127	14	011696	NUCLEOPROTEIN (FRAGMEN
5	53	70.7	265	5	025144	HROX1
6	53	70.7	897	5	017336	LET 858.
7	53	70.7	1238	5	061198	F15E6.6 PROTEIN.
8	52	69.3	372	5	P91143	SIMILAR TO ACETYLTTRANS
9	52	69.3	475	13	093514	AXIAL PROTOCADHERIN (F
10	52	69.3	932	5	001623	SIMILAR TO LIGAND-GATE
11	51	68.0	83	9	048385	ORF83.
12	51	68.0	471	10	023552	HYPOTHETICAL 52.8 KD P
13	51	68.0	582	3	074931	ALTERNATIVE NADH-DEHYD
14	51	68.0	1035	13	057537	NF-PROTOCADHERIN-A.
15	51	68.0	1069	11	088185	BH-PROTOCADHERIN-A.
16	51	68.0	1069	4	060245	PCDH7 (BH-PCDH)A.
17	51	68.0	1072	4	060246	PCDH7 (BH-PCDH)B.
18	51	68.0	1200	4	060247	PCDH7 (BH-PCDH)C.
19	50	66.7	293	2	051095	CONSERVED HYPOTHETICAL
20	50	66.7	444	5	P91141	SIMILAR TO ACETYLTTRANS

Query Match 73.3%; Score 55; DB 3; Length 649;
 Best Local Similarity 54.5%; Pred. No. 2.58e+00;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

DB	Score	Query Match	Length	ID	Description	Pred. No.
234	234	NYLSTARSLEF	244			
2	2	SYLSTSSSLDY	12			
2	2	PRELIMINARY;	349	AA.		
017959;						
017959;						
01-JAN-1998 (TREMREL. 05, CREATED)						
01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)						
01-JAN-1998 (TREMREL. 09, LAST ANNOTATION UPDATE)						
MO1B2.5 PROTEIN.						
MO1B2.5.						
CAENORHABDITIS ELEGANS.						
EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIDA.						
RHABDITIDA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.						
[1]						
SEQUENCE FROM N.A.						
RA LLOYD C.;						
SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.						

[2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE: 94150718.
 RX WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BOWFIELD J., BURTON J., CONNELL M., COPESEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHONKKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THERBY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOLDMAN P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL NATURE 368:32-38(1994).
 DR EMBL: 283116; E1348127; -.
 SO SEQUENCE 349 AA; 40017 MW; 90870FE2 CRC32;

Query Match 72.0%; Score 54; DB 5; Length 349;
 Best Local Similarity 70.0%; Pred. No. 4.07e+00;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 198 YLSSSLDY 207
 QY 3 YLSTSSLDY 12

RESULT 3 PRELIMINARY: PRT: 503 AA.
 ID 026074.
 AC 026074.
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE PROTEIN-EXPORT MEMBRANE PROTEIN (SECD).
 GN HP1550.
 OS HELICOBACTER PYLORI (CAMPILOBACTER PYLORI).
 OC BACTERIA: PROTEOBACTERIA; EPSILON SUBDIVISION; HELICOBACTER GROUP;
 CC HELICOBACTER.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-26695.
 RX MEDLINE: 97394467.
 RA TOUB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,
 RA FELSCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
 RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
 RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,
 RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
 RA BERG D.E., GOCARINE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,
 RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., MATTHEY L., WALLIN E.,
 RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
 RA VENTER J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter
 RT pylori [published erratum appears in Nature 1997 Sep
 RT 25:389(6649):412]."
 RL NATURE 388:539-547(1997).
 DR EMBL: AE000652; G2314730; -.
 DR TIGR: HP1550; -.
 KW HYPOTHETICAL PROTEIN.
 SO SEQUENCE 503 AA; 54247 MW; 9A76592C CRC32;

Query Match 72.0%; Score 54; DB 2; Length 503;
 Best Local Similarity 60.0%; Pred. No. 4.07e+00;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 38 YLSSSLDY 47
 QY 3 YLSTSSLDY 12

RESULT 4 PRELIMINARY: PRT: 127 AA.
 ID 011696
 AC 011696;

DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
 DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE NUCLEOPROTEIN (FRAGMENT)
 OS MEASLES VIRUS (SUBGROUP SCLEPOSE PANENEPHALITIS VIRUS).
 CC VIRUSES: SSRNA NEGATIVE-STRAND VIRUSES: MONONGAVIRALES;
 OC PARAMYXOVIRIDAE: PARAMYXOVIRINAE: MORBILLIVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-92-E;
 RX MEDLINE: 97278133.
 RA YAMAGUCHI S.;
 RT "Identification of three lineages of wild measles virus by nucleotide
 RT sequence analysis of N, P, M, F, and L genes in Japan.";
 RL J. MED. VIROL. 52:113-120(1997).
 DR EMBL: D87487; D1020995; -.
 KW NUCLEOPROTEIN.
 FT NON_TER 1
 SO SEQUENCE 127 AA; 13950 MW; 42D75A2C CRC32;

Query Match 70.7%; Score 53; DB 14; Length 127;
 Best Local Similarity 63.6%; Pred. No. 6.38e+00;
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 66 AYLPSTPLD 76
 QY 1 ASYLSTSSLD 11

RESULT 5 PRELIMINARY: PRT: 265 AA.
 ID 025144.
 AC 025144.
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 GN HROX1.
 DE HROX1.
 OS HALIOTIS RUPESCENS (CALIFORNIA RED ABALONE).
 OC EUKARYOTA: METAZOA: MOLLUSCA: GASTROPODA: PROSOBRANCHIA;
 CC ARCHAEOGASTROPODA; HALIOTIDAE; HALIOTIS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97388321.
 RA DEGNAN B.M., DEGNAN S.M., FENTENAN G., MORSE D.E.;
 RT "A Mox homeobox gene in the gastropod mollusc Haliotis rufescens is
 RT differentially expressed during larval morphogenesis and
 RT metamorphosis".
 RL FEBS LETT. 411:119-122(1997).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL: X75217; G407415; -.
 DR PROSITE: PS00027; HOMEBOX_1; 1.
 DR PFAM: PF00046; homeobox; 1.
 KW HOMEBOX; DNA-BINDING; NUCLEAR PROTEIN.
 SO SEQUENCE 265 AA; 29579 MW; B68A753D CRC32;

Query Match 70.7%; Score 53; DB 5; Length 265;
 Best Local Similarity 75.0%; Pred. No. 6.38e+00;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 81 GSYLSMSSSKDY 92
 QY 1 ASYLSTSSLDY 12

RESULT 6 PRELIMINARY: PRT: 897 AA.
 ID 017336
 AC 017336;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
 DE LET 858.
 GN LET-858.
 OS CAENORHABDITIS ELEGANS.

OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-N2 (BRISTOL);
 RA KELLY W.G., COLES L.H., FIRE A.Z.;
 RL GENETICS 0:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MATTHEWS L.;
 RL SUBMITTED (JUN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U19615; G987227; -;
 DR EMBL: 281525; E1351661; -;
 SQ SEQUENCE 897 AA; 104268 MW; E1E3EA36 CRC32;

Query Match 70.7%; Score 53; DB 5; Length 897;
 Best Local Similarity 70.0%; Pred. No. 6.38e+00;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

491 YLTMOSSLDY 500
 ||:|||||
 3 YLTSSSLDY 12

RESULT 7
 ID 061198 PRELIMINARY; PRT; 1238 AA.
 AC 061198;
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE F15E6.6 PROTEIN.
 GN F15E6.6.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE; 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAYTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HARKINS T., HILLER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
 "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL NATURE 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA MILLER N., STELLIXES L., BRADSHAW H., KEPPLER D.;
 RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA WATERSTON R.;
 RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA WATERSTON R.;
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U88315; G1825777; -;
 DR PFAM: PF00561; abhydrolase; 1.
 KW TRANSFERASE.
 SQ SEQUENCE 372 AA; 42139 MW; 5214F159 CRC32;

Query Match 70.7%; Score 53; DB 5; Length 1238;
 Best Local Similarity 70.0%; Pred. No. 6.38e+00;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Db 702 YLTMOSSLDY 711
 ||:|||||
 1 YLTSSSLDY 12

OY 3 YLTSSSLDY 12

RESULT 8
 ID P91143 PRELIMINARY; PRT; 372 AA.
 AC P91143;
 DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
 DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE SIMILAR TO ACETYLTRANSFERASES.
 GN C37H5.2.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE; 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAYTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HARKINS T., HILLER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
 "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL NATURE 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA DAVIDSON S., GILLAM B.;
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA WATERSTON R.;
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U88315; G1825777; -;
 DR PFAM: PF00561; abhydrolase; 1.
 KW TRANSFERASE.
 SQ SEQUENCE 372 AA; 42139 MW; 5214F159 CRC32;

Query Match 69.3%; Score 52; DB 5; Length 372;
 Best Local Similarity 58.3%; Pred. No. 9.96e+00;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Db 168 GYLTSTYALKY 179
 :|||||:|
 OY 1 ASYLTSSSLDY 12

RESULT 9
 ID 093514 PRELIMINARY; PRT; 475 AA.
 AC 093514;
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE AXIAL PROTOCADHERIN (FRAGMENT).
 GN AXP.
 OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
 OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENODIDINAE; XENOPUS.
 [1]
 RP SEQUENCE FROM N.A.
 RA YAMAMOTO A., DEROBERTIS E.M.;
 RL "Xenopus axial protocadherin.";
 DR EMBL: AF053469; G3596688; -;
 FT NON_TER 1 1

FT NON TER 475 475
SEQUENCE 475 AA: 52268 MW: 2A681544 CRC32;

Query Match 69.3%; Score 52; DB 13; Length 475;
Best Local Similarity 60.0%; Pred. No. 9.96e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 341 FLQTTSLDY 350
: 1:11111
QY 3 YLSTSSLDY 12

RESULT 10
ID 001623; PRELIMINARY; PRT: 932 AA.
AC 001623;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SIMILAR TO LIGAND-GATED IONIC CHANNEL PROTEINS.
ZC196.7.

CAENORHABDITIS ELEGANS.
EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
[1]

SEQUENCE FROM N.A.
STRAIN-BRISTOL N2;
MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BOWFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRATON A., DEAR S., DU Z., DUBBIN R., FAVELLO A., FULTON L.,
RA GARNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SUSTON J.,
RA THERRER-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RT NATURE 368:32-38(1994).
RL [2]

RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA MURRAY J.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: 097007; G1938466; -;
DR PFAM: PF00060; 119_chan; 2.

SO SEQUENCE 932 AA: 106836 MW: 469F21FA CRC32;

Query Match 69.3%; Score 52; DB 5; Length 932;
Best Local Similarity 54.5%; Pred. No. 9.96e+00;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 753 AYLWESTSLEY 763
: 1:11111
QY 2 YLSTSSLDY 12

RESULT 11
ID 048385; PRELIMINARY; PRT: 83 AA.
AC 048385;

DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ORE83.
OS STREPTOCOCCUS THERMOPHILUS BACTERIOPHAGE TP-J34.
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; TAILED PHAGES; SIPHOVIRIDAE.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-TP-J34;
RX MEDLINE: 98122991.
RA NEVE H., ZENZ K.I., DESIERE F., KOCH A., HELLER K.J., BRUSSON H.;
RT "A survey of the lysogenic modules from the temperate Streptococcus
thermophilus bacteriophages TP-J34 and Sf121: implications for the
RT modular theory of phage evolution.";
RL VIROLOGY 241:61-72(1998).
DR EMBL: AF020798; G2897100; -;

SO SEQUENCE 83 AA: 9876 MW: 250B3F9C CRC32;

Query Match 68.0%; Score 51; DB 9; Length 83;
Best Local Similarity 60.0%; Pred. No. 1.55e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 71 YLETSAPLEY 80
: 1:11111
QY 3 YLSTSSLDY 12

RESULT 12
ID 023552; PRELIMINARY; PRT: 471 AA.
AC 023552;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 52.8 KD PROTEIN.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC CAPRIFALES; BRASSICACEAE; ARABIDOPSIS.
RN [1]

RP SEQUENCE FROM N.A.
RA BEVAN M., STEKEMA W., MURPHY G., WAMBUTT R., POHL T., TERRY N.,
RA KREIS M., KAVANAGH T., ENTIAN K.D., RIEGER M., JAMES R.,
RA PUIGDOMENICH P., HATZIOPOULOS P., OBERMAYER B., DUESTERHOFT A.,
RA JONES J., PALME K., ANSGORE W., DELSENY M., BANCROFT I., MEWES H.W.,
RA SCHUELLER C., CHALMARTZIS N.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]

RP SEQUENCE FROM N.A.
RA EU ARABIDOPSIS SEQUENCING PROJECT, ESSA.
SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: 297342; E327038; -;
KM HYPOTHETICAL PROTEIN.

SO SEQUENCE 471 AA: 52785 MW: 4EBA9315 CRC32;

Query Match 68.0%; Score 51; DB 10; Length 471;
Best Local Similarity 63.6%; Pred. No. 1.55e+01;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 19 AGYLTSSDLD 29
: 1:11111111
QY 1 ASYLTSSDLD 11

RESULT 13
ID 074931; PRELIMINARY; PRT: 582 AA.
AC 074931;

DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ALTERNATIVE NADH-DEHYDROGENASE PRECURSOR (EC 1.6.5.3)
DE (NADH DEHYDROGENASE (UBIQUINONE)) (UBIQUINONE REDUCTASE)
DE (TYPE I DEHYDROGENASE) (COMPLEX I DEHYDROGENASE).
GN NDH2.

OS YARROWIA LIPOLYTICA (CANDIDA LIPOLYTICA).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC DIPODASCACEAE; YARROWIA.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-E150;
RA KERSCHER S.J., BRANDT U.;

```

RT "Identification of the YLNDH2 Gene Encoding the Alternative
RL NADH:ubiquinone Oxidoreductase from Yarrowia lipolytica.";
RL SUBMITTED (JUN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
CC -1- COFACTOR: FAD; IRON-SULFUR.
DR EMBL; AJ006852; E1330342; -.
KW SIGNAL; OXIDOREDUCTASE.
FT CHAIN 1 79 POTENTIAL.
SO SEQUENCE 582 AA; 65814 MW; 0460C796 CRC32;

Query Match
Best Local Similarity 54.5%; Score 51; DB 3; Length 582;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 81 TYSIASISG 91
QY 2 YLSTSSSLDY 12

ULT 14 PRELIMINARY; PRT; 1035 AA.
AC 057537;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE NF-PROTOCOLADHERIN.
GN NFPC.
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.
RN [1]
RP SEQUENCE FROM N.A.
RA BRADLEY R.S., ESPSETH A., KINTNER C.;
RL CURR. BIOL. 0:0-0(1998).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AF043643; G2852363; -.
DR PROSITE; PS00232; CADHERIN; 6.
KW CELL ADHESION; GLYCOPROTEIN; TRANSMEMBRANE; CALCIUM-BINDING; REPEAT.
SQ SEQUENCE 1035 AA; 113713 MW; 7E4D3C4E CRC32;

Query Match
Best Local Similarity 60.0%; Score 51; DB 13; Length 1035;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 458 FLHTSAPLDY 467
QY 3 YLSTSSSLDY 12

RESULT 15 PRELIMINARY; PRT; 1069 AA.
ID 088185;
AC 088185;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE BH-PROTOCOLADHERIN-A.
GN PCDH7.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE; 98277460.
RA YOSHIDA K., YOSHITOMO-NAKAGAWA K., SEKI N., SASAKI M., SUGANO S.;
RT "Cloning, expression analysis, and chromosomal localization of
RT BH-protocadherin (PCDH7), a novel member of the cadherin
RT superfamily.";
RL GENOMICS 49:458-461(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA YOSHIDA K., HIDA M., WATANABE M., YAMAGUCHI R., TATEYAMA S.,
RA SUGANO S.;

```

```

RT "cDNA cloning and chromosomal mapping of mouse BH-protocadherin.";
RL DNA SEQ. 0:0-0(1998).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AB006758; D1033562; -.
DR PROSITE; PS00232; CADHERIN; 5.
KW CELL ADHESION; GLYCOPROTEIN; TRANSMEMBRANE; CALCIUM-BINDING; REPEAT.
SQ SEQUENCE 1069 AA; 116314 MW; 0F3F60C6 CRC32;

Query Match
Best Local Similarity 60.0%; Score 51; DB 11; Length 1069;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 486 FLHTSAPLDY 495
QY 3 YLSTSSSLDY 12

Search completed: Thu Sep 2 12:32:56 1999
Job time : 50 secs.

```

THIS PAGE BLANK (USPTO)

CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 12 AA;

Query Match 100.0%; Score 69; DB 27; Length 12;
 Best Local Similarity 100.0%; Pred. No. 7.15e-01;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 aaylstssld 12
 |||||
 QY 1 ASYLSTSSSLDD 12

RESULT 2
 W27586 standard; peptide: 12 AA.

AC W27586;
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KM Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KM heavy chain; complementarity determining region 3; inhibition;
 KM treatment; sepsis; disease; autoimmune disease; infectious disease;
 KM malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KM keloid formation; scar tissue formation; pyrexia; HUVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.

OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997; U02219.
 PF 10-FEB-1997; U02219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRM, Kaymakalan Z, Labkovsky B,
 PI Manovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR WPI: 97-415302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer

Claim 20; Page 72; 102pp; English.
 The present sequence is a novel anti-human tumour necrosis
 factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 12 AA;

Query Match 95.7%; Score 66; DB 27; Length 12;
 Best Local Similarity 91.7%; Pred. No. 1.59e-00;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Db 1 aaylstssld 12
 |||||
 QY 1 ASYLSTSSSLDD 12

RESULT 3
 W27588 standard; peptide: 12 AA.

ID W27588;
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KM Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KM heavy chain; complementarity determining region 3; inhibition;
 KM treatment; sepsis; disease; autoimmune disease; infectious disease;
 KM malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KM keloid formation; scar tissue formation; pyrexia; HUVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.

OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997; U02219.
 PF 10-FEB-1997; U02219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRM, Kaymakalan Z, Labkovsky B,
 PI Manovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR WPI: 97-415302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer

Claim 20; Page 73; 102pp; English.
 The present sequence is a novel anti-human tumour necrosis
 factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 12 AA;

Query Match 91.3%; Score 63; DB 27; Length 12;
 Best Local Similarity 100.0%; Pred. No. 3.52e+00;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 aaylstssld 11
 |||||
 QY 1 ASYLSTSSSLDD 11

RESULT 4
 W27587 standard; peptide: 12 AA.

ID W27587;
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KM Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;

KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disease; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PR 10-FEB-1997: US-02219.
 PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakçalan Z, Labkovsky B,
 PI Menkovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WPI; 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20, Page 73; 102pp: English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:
 SQ

Query Match 91.3% Score 63; DB 27; Length 12;
 Best Local Similarity 100.0%; Pred. No. 3.52e+00;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 asy1stssld 11
 |||||
 QY 1 ASY1STSSLD 11

RESULT 5
 ID W27593 standard; peptide: 12 AA.
 AC W27593;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disease; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PR 10-FEB-1997: US-02219.
 PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakçalan Z, Labkovsky B,
 PI Menkovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WPI; 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20, Page 75; 102pp: English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:
 SQ

PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakçalan Z, Labkovsky B,
 PI Menkovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WPI; 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20, Page 75; 102pp: English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:
 SQ

Query Match 87.0% Score 60; DB 27; Length 12;
 Best Local Similarity 81.8%; Pred. No. 7.70e+00;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 asy1stssld 11
 |||||
 QY 1 ASY1STSSLD 11

RESULT 6
 ID W27594 standard; peptide: 12 AA.
 AC W27594;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disease; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PR 10-FEB-1997: US-02219.
 PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakçalan Z, Labkovsky B,
 PI Menkovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WPI; 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Disclosure; Page 75; 102pp: English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC Sequence 12 AA:
 SQ

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis;
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:

Query Match 87.0%; Score 60; DB 27; Length 12;
 Best Local Similarity 81.8%; Pred. No. 7.70e+00;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 2 sylstassldn 12
 |||||
 0y 2 SYLSTSSSLD 12

RESULT 7
 ID W27563 standard; peptide: 12 AA.
 AC W27563.

DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.

KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.

OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc_difference 12 /label: Tyr, Asn

PN W09729131-A1.
 14-AUG-1997.
 10-FEB-1997; U02219.

PR 25-NOV-1996; US-031476.
 09-FEB-1996; US-599226.

PA (BADI) BASF AG.

PI Allen DJ, Hoogenboom HRM, Kaymakcalan Z, Labkovsky B,
 PI Markovitch JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Saifeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
 DR WPI: 97-415302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer

PS Claim 9; Page 65; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis;
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,

CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:

Query Match 82.6%; Score 57; DB 27; Length 12;
 Best Local Similarity 90.0%; Pred. No. 1.67e+01;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 2 sylstassld 11
 |||||
 0y 2 SYLSTSSSLD 11

RESULT 8
 ID W27592 standard; peptide: 12 AA.
 AC W27592.

DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.

KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.

OS Homo sapiens.
 PN W09729131-A1.

PD 14-AUG-1997.
 PF 10-FEB-1997; U02219.

PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.

PA (BADI) BASF AG.

PI Allen DJ, Hoogenboom HRM, Kaymakcalan Z, Labkovsky B,
 PI Markovitch JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Saifeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
 DR WPI: 97-415302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer

PS Claim 20; Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis;
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:

Query Match 82.6%; Score 57; DB 27; Length 12;
 Best Local Similarity 81.8%; Pred. No. 1.67e+01;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 asfistsssl 11
 11:|||||
 QY 1 ASYSTSSSLD 11

RESULT 9
 ID W27591 standard; peptide: 12 AA.
 AC W27591;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KM Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KM heavy chain; complementarity determining region 3; inhibition;
 KM treatment; sepsis; disease; autoimmune disease; infectious disease;
 KM malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KM keloid formation; scar tissue formation; pyrexia; HVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 KM Homo sapiens.
 KM WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,
 PI Mankevich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfield JG, Schoenhaut D, Vaughan TV, White M, Willton AJ;
 DR WPI: 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, infectious diseases, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HVEC).
 CC Sequence 12 AA;
 SQ

Query Match 82.6%; Score 57; DB 27; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.67e+01;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 asfistsssl 10
 11:|||||
 QY 1 ASYSTSSSL 10

RESULT 10
 ID W27569 standard; Protein; 121 AA.
 AC W27569;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain variable region.
 KM Human: tumour necrosis factor-alpha; TNF-alpha; antibody;
 KM heavy chain; variable region; inhibition;
 KM treatment; sepsis; disease; autoimmune disease; infectious disease;

KM malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KM keloid formation; scar tissue formation; pyrexia; HVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 KM Homo sapiens.
 KM WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,
 PI Mankevich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfield JG, Schoenhaut D, Vaughan TV, White M, Willton AJ;
 DR WPI: 97-415302/38.
 DR N-PDB; T88404.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 16; Page 76; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain variable region.
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, infectious diseases, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HVEC).
 CC Sequence 121 AA;
 SQ

Query Match 82.6%; Score 57; DB 27; Length 121;
 Best Local Similarity 90.0%; Pred. No. 1.67e+01;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 100 systsssl 109
 11:|||||
 QY 2 SYSTSSSLD 11

RESULT 11
 ID W27590 standard; peptide: 12 AA.
 AC W27590;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KM Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KM heavy chain; complementarity determining region 3; inhibition;
 KM treatment; sepsis; disease; autoimmune disease; infectious disease;
 KM malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KM keloid formation; scar tissue formation; pyrexia; HVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 KM Homo sapiens.
 KM WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.

PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Markovitch JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR WPI: 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PS TNF alpha activity, e.g. this to treat autoimmune diseases and cancer
 CC Claim 20: Page 74; 102pp: English.
 CC The present sequence is a novel anti-human tumor necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. Rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 12 AA;

Query Match 79.7%; Score 55; DB 27; Length 12;
 Best Local Similarity 90.9%; Pred. No. 2.78e+01;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 asy1stsfld 11
 111111111111
 Qy 1 ASYLSTSSSLD 11

RESULT 12
 ID W23067 standard; Protein: 417 AA.
 AC W23067; 19-FEB-1998 (first entry)
 DE Canine IGE heavy chain constant region (exon 1-4 product).
 KW IGE; Immunoglobulin; antibody; heavy chain constant region;
 KM allergy; hypersensitivity; therapy; dog; antisense;
 OS immunomodulation.
 PS Cants familiaris.
 CC Key Location/Qualifiers
 CC Misc_difference 55
 FT Misc_difference 55 /note: "encoded by ACC"
 FT Misc_difference 56 /note: "encoded by TAC"
 FT Misc_difference 67 /note: "encoded by GCC"
 FT Misc_difference 83 /note: "encoded by NNT"
 FT Misc_difference 174 /note: "encoded by GGN"
 FT Misc_difference 175 /note: "encoded by NNG"
 FT Misc_difference 176 /note: "encoded by TGN"
 FT Misc_difference 203 /note: "encoded by TCC"
 FT Misc_difference 204 /note: "encoded by GAC"
 FT W09730156-A2.
 PN 21-AUG-1997.
 PD 14-FEB-1997; U02322.
 PR 14-FEB-1996; US-601197.
 PA (INDEX-) IDEXX LAB INC.
 PI Harris RA, Werner B, Steifing AE;

DR WPI: 97-425031/39.
 DR N-PSDB; T79278.
 PT Isolated canine IGE heavy chain constant region DNA - useful to
 PT develop products for treatment of canine allergies and for
 PT immunomodulation in dogs
 PS Disclosure: Page 35-39; 59pp: English.
 CC This polypeptide is encoded by exons 1-4 (see T79278) of canine
 CC IGE heavy chain constant region (epsilon) genomic DNA. Another
 CC polypeptide, comprising the exon 5 and 6 product, is given in
 CC W23068. Recombinant peptides encoded by exons 1-6 can be
 CC produced in eukaryotic or prokaryotic cells. Such peptides,
 CC and antibodies raised against them, are used in methods to treat
 CC the manifestation of allergy in dogs, e.g. to treatment Type I
 CC immediate hypersensitivity, and for immunomodulation.
 SQ Sequence 417 AA;

Query Match 72.5%; Score 50; DB 26; Length 417;
 Best Local Similarity 63.6%; Pred. No. 9.68e+01;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 209 say1appapld 219
 111111111111
 Qy 1 ASYLSTSSSLD 11

RESULT 13
 ID R87027 standard; Protein: 314 AA.
 AC R87027.
 DR 18-MAR-1996 (first entry)
 DE T7 gene 10 leader sequence product.
 KW Plasmid PROPE-1b(+); vector; ligand binding domain; epitope mapping;
 KM antigen; Escherichia coli.
 OS Bacteriophage T7.
 PN US5464745-A.
 PD 07-NOV-1995.
 PF 31-MAR-1993; 040753.
 PR 31-MAR-1993; US-040753.
 PA (NOVA-) NOVAGEN INC.
 PI Garber R, Hammer B, Merendorf R, Novy R;
 DR WPI: 95-392610/50.
 DR N-PSDB; T07310.

PT Mapping ligand binding domains, esp. epitope(s), of proteins - by
 PT expressing peptide(s) encoded by random gene fragments and testing
 PT for ligand binding
 PS Disclosure: Column 15-18; 12pp: English.
 CC The phage T7 gene 10 leader sequence product (R87027) is encoded
 CC by prokaryotic expression vector PROPE-1b(+)(T07310). A fusion
 CC esp. antigenic region, of a putative ligand-binding domain (LBD),
 CC of putative LBD-encoding DNA into the vector. The fusion protein
 CC accumulates as inclusion bodies in Escherichia coli host cells and
 CC can be screened for its ability to bind a ligand.
 SQ Sequence 314 AA;

Query Match 71.0%; Score 49; DB 15; Length 314;
 Best Local Similarity 50.0%; Pred. No. 1.24e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

DB 71 aay1apgenld 82
 111111111111
 Qy 1 ASYLSTSSSLD 12

RESULT 14
 ID W06968 standard; Protein: 384 AA.
 AC W06968;
 DR 12-MAR-1997 (first entry)
 DE M. tuberculosis ORF 2 product.
 KW MACE; mycolic acid cyclopropanating enzyme; cyclopropanation; cma;
 KM pathogenic form; mycobacterium; non-pathogenic; screen; inhibitor;
 KW cyclopropane mycolic acid synthase; beta-ketoacyl reductase;
 KM Streptomyces cinamonensis; homology.
 OS Mycobacterium tuberculosis.

PN US573915-A.
 PD 12-NOV-1996.
 PF 01-JUN-1995; 457245.
 PR 01-JUN-1995; US-457245.
 PA (USGO) US GOVERNMENT.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Clifton B. Yuan Y.
 DR WPI; 96-517880/51.
 DR N-PSDB; T46163.

PT Screening assay for anti-mycobacterial agents - based on inhibition
 of mycolic acid cyclopropanation enzyme
 PS Disclosure: Column 35-38; 35pp; English.
 CC This sequence is encoded by ORF 2 of Mycobacterium tuberculosis (see
 T46163). Three ORFs have been found all within a 7.2 kb BamHI fragment
 isolated from M. tuberculosis, and are believed to be related to the
 biosynthesis of mycolic acids. The ORF 2 gene product has homology to
 CC known enzymes involved in the oxidative/reductive interconversions of
 a ketone and an alcohol. It is most homologous (30 percent identity
 over 188 amino acids) to actIII, beta-ketoacyl reductase from
 Streptomyces cinnamonensis which is involved in chain elongation in
 polyketide biosynthesis. Cyclopropanation of mycolic acids distinguishes
 pathogenic forms of mycobacterium from non-pathogenic forms. A method to
 determine the ability of a cpd. to inhibit cyclopropanation of mycolic
 CC acids in pathogenic mycobacterium is claimed.
 SQ Sequence 384 AA;

Query Match 71.0%; Score 49; DB 20; Length 384;
 Best Local Similarity 54.5%; Pred. No. 1.24e+02;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 235 ssy1ptkald 245
 QY 1 ASYLSTSSSLD 11

RESULT 15
 ID W68459 standard: protein; 386 AA.
 AC W68459;
 DT 06-JAN-1999 (first entry)
 DE PRRSV isolate 14/96 nucleocapsid/phage T7 gene 10 fusion protein.
 KM Nucleocapsid: Porcine Reproductive and Respiratory Syndrome Virus; PRRSV;
 KM p1g; serum; RT-PCR; reverse transcription; amplification; fusion protein;
 KM primer: bacteriophage T7; gene 10; E coli; immunoassay; diagnosis.
 OS Chimeric - Porcine reproductive and respiratory syndrome virus.
 OS Chimeric - Bacteriophage T7.
 PN WO9829550-A1.
 PD 03-JUL-1998.
 PD 24-DEC-1997; ES0313.
 PD 30-DEC-1996; ES-002770.

PI (INMU-) IMMUNOLOGIA & GENETICA APLICADA SA.
 PI Casal Alvarez JI, Rodriguez Garcia MJ, Sanz Fernandez A;
 DR WPI; 98-388131/33.
 PT New fusion proteins comprising Porcine Reproductive and Respiratory
 PT Syndrome Virus nucleocapsid - useful as reagents in immuno:diagnosis
 PT of PRRS, are produced recombinantly in bacterial hosts
 PS Claim 5; Page 31-32; 42pp; Spanish.
 CC This sequence represents a fusion protein comprising the nucleocapsid
 CC protein from the Porcine Reproductive and Respiratory Syndrome Virus
 CC (PRRSV) American Isolate Canada 14/96 fused to the gene 10 protein from
 CC bacteriophage T7. The fusion protein is generated by cloning the PRRSV
 CC nucleocapsid (W68457) coding sequence into the bacterial expression
 CC vector pET3x. The fusion protein is produced when the gene is expressed
 CC in E. coli (BL21) plys cells. The nucleocapsid protein, and especially
 CC the fusion protein, is useful as a reagent in immunoassays to diagnose
 CC PRRSV.
 SQ Sequence 386 AA;

Query Match 71.0%; Score 49; DB 36; Length 386;
 Best Local Similarity 50.0%; Pred. No. 1.24e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 71 aaylapgenldd 82
 QY 1 ASYLSTSSSLD 11

QY 1 ASYLSTSSSLD 12
 Search completed: Thu Sep 2 12:35:05 1999
 Job time : 20 secs.

THIS PAGE BLANK (USPTO)

WIDEVIEW (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_PP protein - protein database search, using Smith-Waterman algorithm
on: Thu Sep 2 12:37:02 1999; Maspar time 1.38 Seconds
Tabular output not generated. 88.598 Million cell updates/sec

Title: >US-08-599-226-30
Description: (1-12) from US08599226.pep
Perfect Score: 69
Sequence: 1 ASYLSPSSSLDP 12

Scoring table: PAM 150
Gap 15

Searched: 106580 segs, 10152877 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PCT9_COMB 4:backfile1

Statistics: Mean 15.800; Variance 47.088; scale 0.336

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the total score being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	50	72.5	426	3	PCT-US95-1	Sequence 2, Applicatio	3.45e+01
2	50	72.5	426	1	US-08-336-	Sequence 2, Applicatio	3.45e+01
3	49	71.0	384	1	US-08-040-	Sequence 2, Applicatio	4.46e+01
4	49	71.0	384	1	US-08-457-	Sequence 5, Applicatio	4.46e+01
5	47	68.1	213	2	US-08-761-	Sequence 10, Applicati	7.38e+01
6	47	68.1	213	1	US-08-287-	Sequence 2, Applicatio	7.38e+01
7	47	68.1	213	2	US-08-761-	Sequence 5, Applicatio	7.38e+01
8	47	68.1	213	1	US-08-459-	Sequence 2, Applicatio	7.38e+01
9	47	68.1	213	1	US-08-459-	Sequence 2, Applicatio	7.38e+01
10	47	68.1	213	1	US-08-460-	Sequence 2, Applicatio	7.38e+01
11	47	68.1	213	2	US-08-761-	Sequence 3, Applicatio	7.38e+01
12	47	68.1	213	3	PCT-US93-0	Sequence 2, Applicatio	7.38e+01
13	46	66.7	20	2	US-08-468-	Sequence 2, Applicatio	9.47e+01
14	46	66.7	21	2	US-08-787-	Sequence 33, Applicati	9.47e+01
15	46	66.7	170	2	US-08-327-	Sequence 1, Applicatio	9.47e+01
16	45	65.2	168	4	5194425-4	Patent No. 5194425.	1.21e+02
17	45	65.2	432	2	US-08-705-	Sequence 18, Applicati	1.21e+02
18	44	63.8	209	2	US-08-771-	Sequence 2, Applicatio	1.55e+02
19	44	63.8	464	1	PCT-US81-0	Sequence 16, Applicati	1.55e+02
20	44	63.8	464	3	PCT-US81-0	Sequence 16, Applicati	1.55e+02
21	43	62.3	170	4	5194425-3	Patent No. 5194425.	1.98e+02
22	43	62.3	170	2	US-08-227-	Sequence 1, Applicatio	1.98e+02
23	43	62.3	238	2	US-08-928-	Sequence 5, Applicatio	1.98e+02

24	43	62.3	240	3	PCT-US95-0	Sequence 80, Applicati	1.98e+02
25	43	62.3	240	1	US-08-261-	Sequence 80, Applicati	1.98e+02
26	43	62.3	345	2	US-08-446-	Sequence 40, Applicati	1.98e+02
27	43	62.3	1255	2	US-08-468-	Sequence 68, Applicati	1.98e+02
28	43	62.3	1255	2	US-08-484-	Sequence 8, Applicatio	1.98e+02
29	43	62.3	1255	1	US-08-467-	Sequence 68, Applicati	1.98e+02
30	43	62.3	1255	2	US-08-486-	Sequence 68, Applicati	1.98e+02
31	43	62.3	1255	2	US-08-625-	Sequence 2, Applicatio	1.98e+02
32	43	62.3	1255	2	US-08-356-	Sequence 2, Applicatio	1.98e+02
33	43	62.3	1255	2	US-08-414-	Sequence 68, Applicati	1.98e+02
34	42	60.9	158	2	US-08-828-	Sequence 3, Applicatio	2.53e+02
35	42	60.9	158	2	US-08-828-	Sequence 1, Applicatio	2.53e+02
36	42	60.9	326	4	5395759-2	Patent No. 5395759.	2.53e+02
37	42	60.9	822	2	US-08-222-	Sequence 7, Applicatio	2.53e+02
38	42	60.9	3666	2	US-08-222-	Sequence 12, Applicati	2.53e+02
39	42	60.9	3727	2	US-08-222-	Sequence 27, Applicati	2.53e+02
40	42	60.9	3778	2	US-08-222-	Sequence 2, Applicatio	2.53e+02
41	41	59.4	113	1	US-08-241-	Sequence 10, Applicati	3.21e+02
42	41	59.4	536	1	US-08-401-	Sequence 2, Applicatio	3.21e+02
43	41	59.4	553	3	PCT-US94-0	Sequence 4, Applicatio	3.21e+02
44	41	59.4	554	2	US-08-464-	Sequence 1, Applicatio	3.21e+02
45	41	59.4	879	2	US-08-486-	Sequence 6, Applicatio	3.21e+02

ALIGNMENTS

RESULT 1
ID PCT-US95-13795-2 STANDARD: PRT: 426 AA.

AC xxxxxx

DE Sequence 2, Application PC/TUS9513795

CC Sequence 2, Application PC/TUS9513795

CC GENERAL INFORMATION:

CC APPLICANT: HOLLIS, GREGORY F.

CC APPLICANT: PATEL, MAYUR D.

CC TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULINS

CC NUMBER OF SEQUENCES: 4

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CHRISTINE E. CARTY

CC STREET: 126 E. LINCOLN AVENUE, P.O. BOX 2000

CC CITY: RAHWAY

CC STATE: NEW JERSEY

CC COUNTRY: USA

CC ZIP: 07065-0907

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC OPERATING SYSTEM: IBM PC compatible

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US95/13795

CC FILING DATE:

CC CLASSIFICATION:

CC ATTORNEY/AGENT INFORMATION:

CC NAME: CARTY, CHRISTINE E.

CC REGISTRATION NUMBER: 36,099

CC REFERENCE/DOCKET NUMBER: 19211Y

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (908) 594-6734

CC TELEFAX: (908) 594-4720

CC INFORMATION FOR SEQ ID NO: 2:

CC LENGTH: 426 amino acids

CC TYPE: amino acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

CC SEQUENCE 426 AA; 47234 MW; 1032622 CN;

Query Match 72.5%; Score 50; DB 3; Length 426;
Best Local Similarity 63.6%; Pred. No. 3,45e+01;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 214 TSYLSPSPSLD 224
:||||:|:|
QY 1 ASYLSTSSSLD 11

RESULT 2
ID US-08-336-583-2 STANDARD; PRT; 426 AA.
XX xxxxxx

Sequence 2, Application US/08336583
Patent No. 5629415
GENERAL INFORMATION:
APPLICANT: HOLLIS, GREGORY F.
APPLICANT: PATEL, MAYUR D.
TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN E
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHRISTINE E. CARTY
STREET: 126 E. LINCOLN AVENUE
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,583
FILING DATE: 09-NOV-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: CARTY, CHRISTINE E.
REGISTRATION NUMBER: 36,099
REFERENCE/DOCKET NUMBER: 19211
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-6734
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 426 AA; 47234 MW; 1032622 CN;

Query Match 72.5%; Score 50; DB 1; Length 426;
Best Local Similarity 63.6%; Pred. No. 3,45e+01;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 214 TSYLSPSPSLD 224
:||||:|:|
QY 1 ASYLSTSSSLD 11

RESULT 3
ID US-08-040-753-2 STANDARD; PRT; 314 AA.
XX xxxxxx
AC xxxxxx
XX xxxxxx
DT
XX

DE Sequence 2, Application US/08040753
XX Sequence 2, Application US/08040753
CC Patent No. 5464745
CC GENERAL INFORMATION:
CC APPLICANT: Mierendorf, Robert
CC APPLICANT: Garder, Richard
CC APPLICANT: No. 5464745Y, Robert
CC APPLICANT: Hammer, Beth
CC TITLE OF INVENTION: Protein Ligand Binding
CC TITLE OF INVENTION: Region Mapping System
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Quarles and Brady
CC STREET: 1 South Pinckney St., Box 2113
CC CITY: Madison
CC STATE: WI
CC COUNTRY: USA
CC ZIP: 53701-2113
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/040,753
CC FILING DATE: 19930331
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Seay, Nicholas J
CC REGISTRATION NUMBER: 27386
CC REFERENCE/DOCKET NUMBER: 70-399-9001-1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 608-251-5000
CC TELEFAX: 608-251-9166
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 314 amino acids
CC TYPE: AMINO ACID
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 314 AA; 33533 MW; 495848 CN;

Query Match 71.0%; Score 49; DB 1; Length 314;
Best Local Similarity 50.0%; Pred. No. 4,46e+01;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 71 AAYLAPGENLDD 82
:||||:|:|
QY 1 ASYLSTSSSLD 12

RESULT 4
ID US-08-457-245-5 STANDARD; PRT; 384 AA.
XX xxxxxx

Sequence 5, Application US/08457245
Patent No. 5573915
GENERAL INFORMATION:
APPLICANT: BARRY III, Clifton E.
APPLICANT: YUAN, Ying
TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA INVOLVED
CC TITLE OF INVENTION: IN THE BIOSYNTHESIS OF CYCLOPROPANATED MYCOLIC ACIDS IN
CC TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS
CC NUMBER OF SEQUENCES: 21
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourile and Crew

REFERENCE/DOCKET NUMBER: CGC 1506/CIP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 213 amino acids
TYPE: amino acid
MOLECULE TYPE: Protein
SEQUENCE 213 AA; 23364 MW; 236956 CN;

Query Match 68.1%; Score 47; DB 1; Length 213;
Best Local Similarity 33.3%; Pred. No. 7.38e+01;
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

DB 99 AGYLTGAGLINE 110
1 ASYLTSSSLD 12

RESULT 7
ID US-08-761-258-5 STANDARD; PRT: 213 AA.
AC xxxxxx
DT
XX
DE
XX
Sequence 5, Application US/08761258
Sequence 5, Application US/08761258
Patent No. 5756087
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Hill, Dwight S.
APPLICANT: Lam, Stephen T.
APPLICANT: Gaffney, Thomas D.
APPLICANT: Torfkwitz, Nancy
TITLE OF INVENTION: Genetically Modified Pseudomonas Strains
TITLE OF INVENTION: With Enhanced Biocontrol Activity
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/761,258
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 213 amino acids
TYPE: amino acid
MOLECULE TYPE: Protein
SEQUENCE 213 AA; 23364 MW; 236956 CN;

Query Match 68.1%; Score 47; DB 2; Length 213;
Best Local Similarity 33.3%; Pred. No. 7.38e+01;
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

DB 99 AGYLTGAGLINE 110
1 ASYLTSSSLD 12

RESULT 8
ID US-08-459-701-2 STANDARD; PRT: 213 AA.
AC xxxxxx
DT
XX
DE
XX
Sequence 2, Application US/08459701
Sequence 2, Application US/08459701
Patent No. 5686282
GENERAL INFORMATION:
APPLICANT: Gaffney, Thomas D.
APPLICANT: Lam, Stephen T.
APPLICANT: Ligon, James M.
APPLICANT: Hill, Dwight S.
APPLICANT: Stein, Jeffrey I.
APPLICANT: Howell, Charles R.
APPLICANT: Becker, J. Ole
TITLE OF INVENTION: Gene Activating Element
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,701
FILING DATE: 02-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/287,442
FILING DATE: 08-AUG-1994
APPLICATION NUMBER: US 08/087,636
FILING DATE: 01-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/908,284
FILING DATE: 02-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/570,184
FILING DATE: 08-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 213 amino acids
TYPE: amino acid
MOLECULE TYPE: Protein
SEQUENCE 213 AA; 23364 MW; 236956 CN;

Query Match 68.1%; Score 47; DB 1; Length 213;
Best Local Similarity 33.3%; Pred. No. 7.38e+01;
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Db 99 AGYLTGAGLINE 110
1:1:1:1:1:
OY 1 ASYLTSSSLD 12

RESULT 9
ID US-08-459-174-2 STANDARD: PRT: 213 AA.
XX
AC xxxxxx
DT
DE
XX

Sequence 2, Application US/08459174

CC Sequence 2, Application US/08459174

CC Patent No. 5710031

CC GENERAL INFORMATION:

CC APPLICANT: Gaffney, Thomas D.

CC APPLICANT: Lam, Stephen T.

CC APPLICANT: Ligon, James M.

CC APPLICANT: Hill, Dwight S.

CC APPLICANT: Stein, Jeffrey I.

CC APPLICANT: Howell, Charles R.

CC APPLICANT: Becker, J. Ole

CC TITLE OF INVENTION: Gene Activating Element

CC NUMBER OF SEQUENCES: 9

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CIBA-GEIGY Corporation

CC STREET: 7 Skyline Drive

CC CITY: Hawthorne

CC STATE: New York

CC COUNTRY: USA

CC ZIP: 10532

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/459,174

CC FILING DATE: 02-JUN-1995

CC CLASSIFICATION: 536

CC PRIORITY APPLICATION DATA:

CC APPLICATION NUMBER: US 08/287,442

CC FILING DATE: 08-AUG-1994

CC APPLICATION NUMBER: US 08/087,636

CC FILING DATE: 01-JUL-1993

CC PRIORITY APPLICATION DATA:

CC APPLICATION NUMBER: US 07/908,284

CC FILING DATE: 02-JUL-1992

CC PRIORITY APPLICATION DATA:

CC APPLICATION NUMBER: US 07/570,184

CC FILING DATE: 08-AUG-1990

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Elmer, James Scott

CC REGISTRATION NUMBER: 36,129

CC REFERENCE/DOCKET NUMBER: CGC 1506/CIP4

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 919-541-8614

CC TELEFAX: 919-541-8689

CC INFORMATION FOR SEQ. ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 213 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

CC SEQUENCE 213 AA; 23364 MW; 236956 CN;

Query Match 68.1%; Score 47; DB 1; Length 213;

Best Local Similarity 33.3%; Pred. No. 7.38e+01;
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Db 99 AGYLTGAGLINE 110
1:1:1:1:1:
OY 1 ASYLTSSSLD 12

RESULT 10
ID US-08-460-298-2 STANDARD: PRT: 213 AA.
XX
AC xxxxxx
DT
DE
XX

Sequence 2, Application US/08460298

CC Sequence 2, Application US/08460298

CC Patent No. 5686283

CC GENERAL INFORMATION:

CC APPLICANT: Gaffney, Thomas D.

CC APPLICANT: Lam, Stephen T.

CC APPLICANT: Ligon, James M.

CC APPLICANT: Hill, Dwight S.

CC APPLICANT: Stein, Jeffrey I.

CC APPLICANT: Howell, Charles R.

CC APPLICANT: Becker, J. Ole

CC TITLE OF INVENTION: Gene Activating Element

CC NUMBER OF SEQUENCES: 9

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CIBA-GEIGY Corporation

CC STREET: 7 Skyline Drive

CC CITY: Hawthorne

CC STATE: New York

CC COUNTRY: USA

CC ZIP: 10532

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/460,298

CC FILING DATE: 02-JUN-1995

CC CLASSIFICATION: 536

CC PRIORITY APPLICATION DATA:

CC APPLICATION NUMBER: US 08/287,442

CC FILING DATE: 08-AUG-1994

CC APPLICATION NUMBER: US 08/087,636

CC FILING DATE: 01-JUL-1993

CC PRIORITY APPLICATION DATA:

CC APPLICATION NUMBER: US 07/908,284

CC FILING DATE: 02-JUL-1992

CC PRIORITY APPLICATION DATA:

CC APPLICATION NUMBER: US 07/570,184

CC FILING DATE: 08-AUG-1990

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Elmer, James Scott

CC REGISTRATION NUMBER: 36,129

CC REFERENCE/DOCKET NUMBER: CGC 1506/CIP4

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 919-541-8614

CC TELEFAX: 919-541-8689

CC INFORMATION FOR SEQ. ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 213 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

CC SEQUENCE 213 AA; 23364 MW; 236956 CN;

Query Match 68.1%; Score 47; DB 1; Length 213;

Best Local Similarity 33.3%; Pred. No. 7.38e+01;
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASYLSTSSSLDD 12

RESULT 11
ID US-08-761-258-3 STANDARD: PRT: 213 AA.

AC xxxxxx

DE Sequence 3, Application US/08761258

CC Sequence 3, Application US/08761258
CC Patent No. 5756087

CC GENERAL INFORMATION:

CC APPLICANT: Ligon, James M.

CC APPLICANT: Hill, Dwight S.

CC APPLICANT: Lam, Stephen T.

CC APPLICANT: Gaffney, Thomas D.

CC APPLICANT: Tokewitz, Nancy

CC TITLE OF INVENTION: Genetically Modified Pseudomonas Strains

CC TITLE OF INVENTION: with Enhanced Biocontrol Activity

CC NUMBER OF SEQUENCES: 11

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Ciba-Geigy Corporation

CC STREET: 520 White Plains Road, P.O. Box 2005

CC CITY: Tarrytown

CC STATE: NY

CC COUNTRY: USA

CC ZIP: 10591

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/761,258

CC FILING DATE:

CC CLASSIFICATION: 424

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Meigs, J. Timothy

CC REGISTRATION NUMBER: 38,241

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (919) 541-8587

CC TELEFAX: (919) 541-8689

CC INFORMATION FOR SEQ ID NO: 3:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 213 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: Protein

CC SEQUENCE 213 AA; 23364 MW; 236956 CN;

CC Query Match 68.1%; Score 47; DB 2; Length 213;

CC Best Local Similarity 33.3%; Pred. No. 7.38e+01;

CC Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Db 99 AGYLTGAGLNE 110

QY 1 ASYLSTSSSLDD 12

RESULT 12
ID PCR-US93-06300A-2 STANDARD: PRT: 213 AA.

AC xxxxxx

DE Sequence 2, Application PC/TUS9306300A

CC Sequence 2, Application PC/TUS9306300A

CC GENERAL INFORMATION:

CC APPLICANT: CIBA-GEIGY AG

CC APPLICANT: Klybeckstrasse 141

CC APPLICANT: 4002 Basle

CC APPLICANT: Switzerland

CC APPLICANT: 125 Tradescant Road

CC APPLICANT: Chapel Hill, NC 27514

CC APPLICANT: USA

CC APPLICANT: 8900 Jeanew Court

CC APPLICANT: Raleigh, NC 27613

CC APPLICANT: USA

CC APPLICANT: Hill, Dwight Steven

CC APPLICANT: Cary, NC 27511

CC APPLICANT: USA

CC APPLICANT: Stehn, Jeffrey I.

CC APPLICANT: 3725 Surry Trail

CC APPLICANT: Hillsborough, NC 27278

CC APPLICANT: USA

CC APPLICANT: Howell, Charles R.

CC APPLICANT: 805 Avondale

CC APPLICANT: Bryan, TX 77802

CC APPLICANT: USA

CC APPLICANT: Becker, J. Ole

CC APPLICANT: 6164 Osevego

CC APPLICANT: Riverside, CA 92506

CC APPLICANT: USA

CC APPLICANT: Ligon, James M.

CC APPLICANT: 120 Marquette Drive

CC APPLICANT: Cary, NC 27513

CC APPLICANT: USA

CC TITLE OF INVENTION: Gene Activating Element

CC NUMBER OF SEQUENCES: 7

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CIBA-GEIGY Corporation

CC STREET: 7 Skyline Drive

CC CITY: Hawthorne

CC STATE: New York

CC COUNTRY: USA

CC ZIP: 10532

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US93/06300A

CC FILING DATE: 02-JUL-1993

CC CLASSIFICATION:

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 07/908,284

CC FILING DATE: 02-JUL-1992

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Spruill, W. Murray

CC REGISTRATION NUMBER: 32,943

CC REFERENCE/DOCKET NUMBER: S-18210/A/CGC1506/PC

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (919)541-8615

CC TELEFAX: (919)541-8689

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 213 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

CC SEQUENCE 213 AA; 23364 MW; 236956 CN;

CC Query Match 68.1%; Score 47; DB 3; Length 213;

CC Best Local Similarity 33.3%; Pred. No. 7.38e+01;

CC Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Db 99 AGYLTGAGLNE 110

QY 1 ASYLSTSSSLDD 12

CC NUMBER OF SEQUENCES: 1
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Venable, Baetjer, Howard & Civiletti
CC STREET: 1201 New York Avenue, N.W., Suite 1000
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA
CC ZIP: 20005
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/327,357A
CC FILING DATE: 21-OCT-1994
CC CLASSIFICATION: 514
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/798,099
CC FILING DATE: 27-NOV-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: CA 2,053,799-0
CC FILING DATE: 22-OCT-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Innen, Jeffrey L.
CC REGISTRATION NUMBER: 28,957
CC REFERENCE/DOCKET NUMBER: 27052-115469
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 202-962-4810
CC TELEFAX: 202-962-8300
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 170 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC ORIGINAL SOURCE:
CC ORGANISM: Homo sapiens
CC IMMEDIATE SOURCE:
CC CLONE: human myelin basic protein
CC SEQUENCE 170 AA; 18459 MW; 143992 CN;
SQ

Query Match 66.7%; Score 46; DB 2; Length 170;
Best Local Similarity 45.5%; Pred. No. 9.47e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

12 SKYLATASMD 22
: ||:|:|:|
1 ASYLSTSSSLD 11

Search completed: Thu Sep 2 12:37:11 1999
Job time : 9 secs.

Stiekema, W.; Drost, L.; Ridley, P.; Hudson, S.A.; Patel, K.; Murphy, G.; Piffanelli, P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terryn, N.; Gielen, J.; Villarroel, R.; De Clerck, R.; Van Montagu, M.; Lecharny, A.; Auborg, S.; Gy, I.; Kreiss, M.; Lao, N.; Kavanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaefer, M.; Funk, B.; Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech, P.; Douka, A.; Vouklatou, E.; Milioni, D.; Hatzopoulos, P.; Piravandi, E.; Obermaier, B.; Hilbert, H.; Duesterhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reckman, S.; Ansoorge, W.; Cooke, R.; Berger, C.; Delsen, M.; Voet, M.; Voelckert, G.; Mewes, H.W.; Klosterman, S.; Schueller, C.; Chaiwatiz, N.
Schueller, C.; Chaiwatiz, N.
Nature (1998) 391:485-488
Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*.
#journal #title Nature (1998) 391:485-488
#cross-references MUID:98121113
#accession C71439
#status preliminary; nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-471 #label BVY
#cross-references GB:297342; NID:g2245031; PID:e327038; PID:g2245065
GENETICS #map_position 4COP9-4G3845
SUMMARY #length 471 #molecular-weight 52785 #checksum 7455
Query Match 73.9%; Score 51; DB 2; Length 471;
Best Local Similarity 63.6%; Pred. No. 3.11e+00;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 19 AGYTRSSDLD 29
1:||||| 11
QY 1 ASYLTSSDLD 11
RESULT 3
ENTRY D70108 #type complete
TITLE conserved hypothetical protein BB0068 - Lyme disease
ORGANISM #formal_name Borrelia burgdorferi #common_name Lyme disease
#spirochete
DATE 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change
05-Jun-1998
D70108
A70100
ACCESSIONS Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Claydon, R.; Lathigra, R.; White, O.; Ketchum, K.A.; Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb, J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.; Kierlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman, J.; Uitterback, T.; Watthey, L.; McDonald, L.; Artach, P.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.
#journal #title Nature (1997) 390:580-586
#cross-references MUID:98065943
#accession D70108
#status preliminary; nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-293 #label KTF
#cross-references GB:AE001120; GB:AE000783; NID:g2687951; PID:g2687956; TIGR:BB0068
#experimental_source strain B31
SUMMARY #length 293 #molecular-weight 33278 #checksum 5223
Query Match 72.5%; Score 50; DB 2; Length 293;
Best Local Similarity 60.0%; Pred. No. 5.12e+00;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 198 AYLTSPNSLE 207
:|||||:
QY 2 SYLTSSSDLD 11
RESULT 4
ENTRY VABPA7 #type complete
TITLE major capsid protein 10A - phage T7
ORGANISM #formal_name phage T7
#note host Escherichia coli
DATE 13-Jun-1983 #sequence_revision 30-Sep-1990 #text_change
26-Feb-1999
ACCESSIONS A04344; S42325
REFERENCE A94615
#authors Dunn, J.J.; Thompson, K.
#submission submitted to the Nucleic Acid Sequence Database, September 1982
#accession A04344
#molecule_type DNA
#residues 1-345 #label DUN
REFERENCE S42283
#authors Dunn, J.J.; Studier, F.W.
#journal J. Mol. Biol. (1983) 166:477-535
#title Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements.
#cross-references MUID:83241725
#accession S42325
#molecule_type DNA
#residues 1-345 #label DUN
#cross-references EMBL:V01146; NID:g431187; PID:g15604
#note the authors did not translate the codon for residue 1
#gene 10A
#map_position 57.51-60.49
CLASSIFICATION #superfamily phage T7 major capsid protein 10A
KEYWORDS capsid protein
SUMMARY #length 345 #molecular-weight 36545 #checksum 8044
Query Match 71.0%; Score 49; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 8.36e+00;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Db 71 MAYIAPGENDLD 82
1:||||| 11
QY 1 ASYLTSSSDLD 12
RESULT 5
ENTRY G71906 #type complete
TITLE probable transcription regulator - Helicobacter pylori
ORGANISM #formal_name Helicobacter pylori
#strain J99
DATE 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
12-Feb-1999
ACCESSIONS G71906
REFERENCE A71800
#authors Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Noonan, B.; Guild, B.C.; deGonge, B.L.; Carmel, G.; Tumblino, P.J.; Caruso, A.; Uria-Mckelsen, M.; Mills, S.D.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Voyls, G.F.; Trust, T.J.
#journal #title Nature (1999) 397:176-180
#cross-references MUID:99120557
#accession G71906
#status preliminary
#molecule_type DNA
#residues 1-381 #label ARN
#cross-references GB:AE001496; GB:AE001499; NID:g4155191; PID:g4155196

RESULT 9
ENTRY S25204 #type complete
TITLE srnx protein - Streptomyces ambofaciens
ORGANISM #formal.name Streptomyces ambofaciens
DATE 28-May-1993 #sequence_revision 28-May-1993 #text_change 01-May-1998

ACCESSIONS
REFERENCE S25204; S21599
#journal S25202
#authors Geistlich, M.; Losick, R.; Turner, J.R.; Rao, R.N.
#title Mol. Microbiol. (1992) 6:2019-2029
Characterization of a novel regulatory gene governing the expression of a polyketide synthase gene in Streptomyces ambofaciens.

#accession S25204
#molecule_type DNA
#residues 1-239 ##label GEI
#cross-references EMBL:X63451; NID:g46699; PID:g46702

GENETICS
#gene srnx
CLASSIFICATION #superfamily bloc homology
FEATURE 39-139 #domain bloc homology #label BIOC
SUMMARY #length 239 #molecular_weight 26493 #checksum 3795

Query Match 69.6%; Score 48; DB 2; Length 239;
Best Local Similarity 45.5%; Pred. No. 1.35e+01;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 110 GYLETADLED 120
:111::111
QY 2 SYLSTSSLD 12

RESULT 10
ENTRY CA9904 #type complete
TITLE L-lactate dehydrogenase (EC 1.1.1.27), FMN-dependent - Escherichia coli
ALTERNATE_NAMES lcid protein
ORGANISM #formal.name Escherichia coli
DATE 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 17-Mar-1999

ACCESSIONS
REFERENCE CA9904; G65160; S47826
#journal CA9904
#authors Dong, J.M.; Taylor, J.S.; Latour, D.J.; Tuchi, S.; Lin, E.C.C.
#title J. Bacteriol. (1993) 175:6671-6678
Three overlapping lct genes involved in L-lactate utilization by Escherichia coli.

#cross-references MUID:94012541
#accession CA9904
#status Preliminary
#molecule_type DNA
#residues 1-396 ##label DON
#cross-references GB:L13970; NID:g404692; PID:g404695

REFERENCE A64720
#journal A64720
#authors Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.
#title Science (1997) 277:1453-1462
The complete genome sequence of Escherichia coli K-12.
#cross-references MUID:97426617
#accession G65160
#status Nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-396 ##label BLAT
#cross-references GB:AE000438; GB:U00096; NID:g2367251; PID:g1790033; UMG:P:3605

REFERENCE S47826
#journal S47826
#authors Plunkett, G.
#submission submitted to the EMBL Data Library, March 1994

#accession S47826
#status Preliminary
#molecule_type DNA
#residues 1-118, 'X', 120-396 ##label PLU
#cross-references EMBL:U00039; NID:g466582; PID:g466743

GENETICS
#gene lcid
CLASSIFICATION #superfamily (S)-2-hydroxy-acid oxidase; (S)-2-hydroxy-acid oxidase homology
KEYWORDS FMN; oxidoreductase
FEATURE 1-321
2/5 #domain (S)-2-hydroxy-acid oxidase homology #label 2HV
SUMMARY #length 396 #molecular_weight 42728 #checksum 9705

Query Match 69.6%; Score 48; DB 2; Length 396;
Best Local Similarity 33.3%; Pred. No. 1.35e+01;
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Db 208 SAVLGRPTLED 219
:111::111
QY 1 ASYLSTSSLD 12

RESULT 11
ENTRY A55033 #type complete
TITLE keratin 12 - mouse
ORGANISM #formal.name Mus musculus #common.name house mouse
DATE 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 17-Mar-1999

ACCESSIONS
REFERENCE A55033
#authors Liu, C.Y.; Zhu, G.; Converse, R.; Kao, C.W.C.; Nakamura, H.; Tseng, S.C.G.; Mui, M.M.; Seiver, J.; Justice, M.J.; Stech, M.E.; Hansen, G.M.; Kao, W.W.Y.
#journal J. Biol. Chem. (1994) 269:24627-24636
#title Characterization and chromosomal localization of the cornea-specific murine keratin gene Krt1.12.
#cross-references MUID:95014223
#accession A55033
#status Preliminary
#molecule_type DNA
#residues 1-483 ##label LIU
#cross-references GB:U08095; NID:g565659; PID:g565660
#note authors translated the codon ATC for residue 225 as Thr, and GCG for residue 388 as Arg

CLASSIFICATION #superfamily cytoskeletal keratin
SUMMARY #length 483 #molecular_weight 52010 #checksum 7

Query Match 69.6%; Score 48; DB 2; Length 483;
Best Local Similarity 50.0%; Pred. No. 1.35e+01;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 131 ASYLGRVSLLE 142
:111::111
QY 1 ASYLSTSSLD 12

RESULT 12
ENTRY A41905 #type complete
TITLE ferric vibriobactin receptor Viua - Vibrio cholerae
ALTERNATE_NAMES iron-siderophore complex receptor Viua
ORGANISM #formal.name Vibrio cholerae
DATE 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Mar-1998

ACCESSIONS
REFERENCE A41905
#journal A41905
#authors Butterton, J.R.; Stoeckner, J.A.; Payne, S.M.; Calderwood, S.B.
#title J. Bacteriol. (1992) 174:3729-3738
Cloning, sequencing, and transcriptional regulation of viua, the gene encoding the ferric vibriobactin receptor of Vibrio cholerae.

#cross-references MUID:92276356
#contents 0395
#accession A41905
#status Preliminary
#molecule_type DNA
#residues 1-687 ##label BUT
#cross-references GB:U11759; GB:M90461; NID:g517206; PID:g531822
##note sequence extracted from NCBI backbone (NCBIN:104700, NCBI:P:104701)

SUMMARY #length 687 #molecular-weight 76412 #checksum 2802

Query Match 69.6%; Score 48; DB 2; Length 687;
Best Local Similarity 60.0%; Pred. No. 1.35e+01;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 662 YLSTNNLTDQ 671
|||:||||:
QY 3 YLSTSSLD 12

ENTRY 13
TITLE S55180 #type complete
phospholipase D homolog YUL132w - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES protein j0678
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 05-Jun-1998

ACCESSIONS S55180; S56914; S71665
REFERENCE S55159
#authors Katsoulou, C.; Tzermia, M.; Alexandraki, D.
#submission submitted to the EMBL Data Library, May 1995
#description The complete sequence of a 40.7 kb segment located on the left arm of yeast chromosome X identified 13 known genes and revealed 13 new open reading frames including homologues to other yeast hypothetical proteins.

#accession S55180
##molecule_type DNA
##residues 1-750 ##label KAT
#cross-references EMBL:X87371; NID:g854542; PID:g854564

REFERENCE S56912
#authors Katsoulou, C.; Tzermia, M.; Alexandraki, D.
#submission submitted to the Protein Sequence Database, September 1995
#accession S56914
##molecule_type DNA
##residues 1-750 ##label KAW
#cross-references EMBL:Z49407; NID:g1008335; PID:g1008336; MIPS:YUL132w

REFERENCE S71643
#authors Katsoulou, C.; Tzermia, M.; Tavernarakis, N.; Alexandraki, D.
#journal Yeast (1996) 12:787-797
#title Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X reveals 14 known genes and 13 new open reading frames including homologues of genes clustered on the right arm of chromosome XI.

#accession S71665
##status nucleic acid sequence not shown
##molecule_type DNA
##residues 1-750 ##label KAF
#cross-references EMBL:X87371; NID:g854542; PID:g854564

GENETICS
#map_position 10L
SUMMARY #length 750 #molecular-weight 84466 #checksum 4191

Query Match 69.6%; Score 48; DB 2; Length 750;
Best Local Similarity 50.0%; Pred. No. 1.35e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 354 YLAVSAPLED 363
||:|:|:|:
QY 3 YLSTSSLD 12

RESULT 14

ENTRY A38222 #type complete
TITLE regulatory protein gacA - Pseudomonas fluorescens
ORGANISM #formal_name Pseudomonas fluorescens
DATE 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Dec-1997

ACCESSIONS A38222
REFERENCE A38222
#authors Laville, J.; Voisard, C.; Keel, C.; Maurhofer, M.; Defago, G.; Haas, D.
#journal Proc. Natl. Acad. Sci. U.S.A. (1992) 89:1562-1566
#title Global control in Pseudomonas fluorescens mediating antibiotic synthesis and suppression of black root rot of tobacco.

#cross-references MUID:92179223
#contents CHA0
#accession A38222
##molecule_type DNA
##residues 1-213 ##label LAV
##note this sequence is inconsistent with the nucleotide translation
##note sequence extracted from NCBI backbone (NCBIN:85369, NCBI:P:85373)

GENETICS
#gene gacA
CLASSIFICATION #superfamily regulatory protein comA; response regulator homology
KEYWORDS phosphoprotein
FEATURE 4-115
54
#domain response regulator homology #label RRH\
#binding site phosphate (Asp) (covalent) #status Predicted

SUMMARY #length 213 #molecular-weight 23663 #checksum 7318

Query Match 68.1%; Score 47; DB 2; Length 213;
Best Local Similarity 33.3%; Pred. No. 2.18e+01;
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Db 99 AGYLRKAGLINE 110
|:|:|:|:|:
QY 1 ASYLTSSLD 12

RESULT 15
ENTRY S42424
TITLE hypothetical protein y - Mycobacterium smegmatis
ORGANISM #formal_name Mycobacterium smegmatis
DATE 07-May-1998 #sequence_revision 15-May-1998 #text_change 15-May-1998

ACCESSIONS S42424; S31804
REFERENCE S42421
#authors Cirillo, J.D.; Weisbrod, T.R.; Pascopella, L.; Bloom, B.R.; Jacobs Jr., W.R.
#journal Mol. Microbiol. (1994) 11:629-639
#title Isolation and characterization of the aspartokinase and aspartate semialdehyde dehydrogenase operon from mycobacteria.

#accession S42424
##molecule_type DNA
##residues 1-333 ##label CIR
#cross-references EMBL:Z17372; NID:g44506; PID:g581353
##note the authors translated the initiation codon GTG for residue 1 as Val

GENETICS
#start_codon GTG
SUMMARY #length 333 #molecular-weight 35881 #checksum 4453

Query Match 68.1%; Score 47; DB 2; Length 333;
Best Local Similarity 41.7%; Pred. No. 2.18e+01;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 290 GGYISPASTD 301
::|:|:|:|:
QY 1 ASYLTSSLD 12

Search completed: Thu Sep 2 12:35:41 1999
Job time : 17 secs.

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE NUCLEOPROTEIN (FRAGMENT).
 OS MEASLES VIRUS (SUBACUTE SCLEROSE PANENCEPHALITIS VIRUS).
 OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; MONONEGAVIRALES;
 CC PARAMYXOVIRIDAE; PARAMYXOVIRINAE; MORBILLIVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-92-E;
 RX MEDLINE: 97278133.
 RA YAMAGUCHI S.;
 RT "Identification of three lineages of wild measles virus by nucleotide
 sequence analysis of N, P, M, F, and L genes in Japan.";
 RL J. MED. VIROL. 52:113-120(1997).
 DR EMBL: D87487; D1020995; -.
 KW NUCLEOPROTEIN.
 FT NON_TER
 SQ SEQUENCE 127 AA; 13950 MW; 42D75A2C CRC32;
 Query Match 76.8%; Score 53; DB 14; Length 127;
 Best Local Similarity 63.6%; Pred. No. 9.36e-01;
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 Db 66 AAYLSTPLD 76
 Qy 1 AYLSTSSLD 11
 RESULT 3
 ID 023552 PRELIMINARY; PRT; 471 AA.
 AC 023552;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 52.8 KD PROTEIN.
 OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
 CC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
 OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
 CC CAPRALES; BRASSICACEAE; ARABIDOPSIS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BEVAN M., STIEKEMA W., MURPHY G., WAMBUTT R., POHL T., TERRYN N.,
 RA KREIS M., KAVANAGH T., ENTIAN K.D., RIEGER M., JAMES R.,
 RA PUIGOMENEGH P., HATZOPOLLOS P., OBERMAIER B., DUESTERHOFF A.,
 RA JONES J., PALME K., ANSOMGE W., DEJSENY M., BANCROFT I., MEMES H.W.,
 RA SCHUELLER C., CHALMATEIS N.;
 RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU ARABIDOPSIS SEQUENCING PROJECT, ESSA;
 DR SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: Z97342; E327038; -.
 KW HYPOTHETICAL PROTEIN.
 FT NON_TER
 SQ SEQUENCE 471 AA; 52785 MW; 4EBA8315 CRC32;
 Query Match 73.9%; Score 51; DB 10; Length 471;
 Best Local Similarity 63.6%; Pred. No. 2.69e+00;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Db 19 AGYLTSSDLD 29
 Qy 1 AYLSTSSLD 11
 RESULT 4
 ID 051095 PRELIMINARY; PRT; 293 AA.
 AC 051095;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE CONSERVED HYPOTHETICAL PROTEIN.
 GN BB0068.
 OS BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).
 CC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35210 / B31;
 RX MEDLINE: 98065943.
 RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
 RA LATHIGER R., WHITE O., KETCHUM K.A., DOSON R., HICKER E.K., GRINN M.,
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
 RA VAN VUUT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
 RA UTTERBACK T., MATTHEY L., MCDONALD L., ARTICH P., BOWMAN C.,
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
 RA SMITH H.O., VENTER J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
 burgdorferi.";
 RL NATURE 390:580-586(1997).
 DR EMBL: AE001120; G2687956; -.
 DR TIGR: BB0068; -.
 SQ SEQUENCE 293 AA; 33278 MW; 3FB9B9E2 CRC32;
 Query Match 72.5%; Score 50; DB 2; Length 293;
 Best Local Similarity 60.0%; Pred. No. 4.51e+00;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 Db 198 AYLTSPNSLE 207
 Qy 2 SYLSTSSLD 11
 RESULT 5
 ID 081650 PRELIMINARY; PRT; 188 AA.
 AC 081650;
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE THYMIDYLATE KINASE (FRAGMENT).
 OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
 CC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
 OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
 CC CAPRALES; BRASSICACEAE; ARABIDOPSIS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA ULLAH H., ROBERTSON N., FITES R.C.;
 RT "Plant thymidylate kinase mRNA.";
 RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF081570; G3493131; -.
 FT NON_TER
 SQ SEQUENCE 188 AA; 21248 MW; C306E7B8 CRC32;
 Query Match 71.0%; Score 49; DB 10; Length 188;
 Best Local Similarity 58.3%; Pred. No. 7.51e+00;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Db 43 SAYLSNKSOLD 54
 Qy 1 AYLSTSSLD 12
 RESULT 6
 ID 011197 PRELIMINARY; PRT; 384 AA.
 AC 011197;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 41.9 KD PROTEIN IN CNA1 3'REGION (ORF2).
 OS MYCOBACTERIUM TUBERCULOSIS.
 CC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
 OC ACTINOMYCETALES; CORYNEBACTERIINAE; MYCOBACTERIACEAE; MYCOBACTERIUM.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RA;
 RX MEDLINE: 95327699.
 RA YUAN Y., LEE R.E., BESRA G.S., BELISLE J.T., BARRY C.E. III;
 RT "Identification of a gene involved in the biosynthesis of

RT cyclopropanated mycolic acids in Mycobacterium tuberculosis.";
RL PROC. NATL. ACADE. SCI. U.S.A. 92:6630-6634(1995).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES.
CC FAMILY (SDS)
DR EMBL: U27357; G886103; -
PFAM: PF00106; adh_short; 1.
KW HYPOTHETICAL PROTEIN: OXIDOREDUCTASE.
SQ SEQUENCE 384 AA; 41995 MW; 9FF8465A CRC32;

Query Match 71.0%; Score 49; DB 2; Length 384;
Best Local Similarity 54.5%; Pred. No. 7.51e+00;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 235 SSYLPTRKALD 245
:||||:|
QY 1 ASYLSTSSSLD 11

RESULT 7
008305; PRELIMINARY; PRT; 437 AA.
008305;

DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE INULIN FRUCTOTRANSFERASE (EC 2.4.1.93).
GN IFT.
OS ARTHROBACTER SP.
OC BACTRRIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIAE.
OC ACTINOMYCETALES; MICROCOCCINEAE; MICROCOCCACEAE; ARTHROBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97179800.
RA SAKURAI H., YOKOTA A., TOMITA F.;
RT "Molecular cloning of an inulin fructotransferase (depolymerizing)
RT gene from Arthrobacter sp. H65-7 and its expression in Escherichia
RT coli.";
RL BIOSCI. BIOTECHNOL. BIOCHEM. 61:87-92(1997).
DR EMBL: D84399; D1019710; -
KW TRANSFERASE; GLYCOSYLTRANSFERASE.
SQ SEQUENCE 437 AA; 46475 MW; DC74B68C CRC32;

Query Match 71.0%; Score 49; DB 2; Length 437;
Best Local Similarity 50.0%; Pred. No. 7.51e+00;
Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Db 331 PPLGTSNGSLD 342
:||||:|
QY 1 ASYLSTSSSLD 12

RESULT 8
ID 050417; PRELIMINARY; PRT; 650 AA.
AC 050417;

DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MULTI-FUNCTIONAL ENZYME.
GN MTV004.49.
OS MYCOBACTERIUM TUBERCULOSIS.
OC BACTRRIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIAE;
OC ACTINOMYCETALES; CORNYBACTERIAE; MYCOBACTERIAE; MYCOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA OLIVER K., SKELTON J., BADCOCK K., CHURCHER C.M., HARRIS D.;
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN-H37RV;
RX MEDLINE: 96181548.
RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,
RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RT "An integrated map of the genome of the tubercle bacillus,
RT Mycobacterium tuberculosis H37RV, and comparison with Mycobacterium
RT leprae";
RL PROC. NATL. ACADE. SCI. U.S.A. 93:3132-3137(1996).
DR EMBL: AL009198; E1202309; -
SQ SEQUENCE 650 AA; 70939 MW; 1E552E85 CRC32;

Query Match 71.0%; Score 49; DB 2; Length 650;
Best Local Similarity 54.5%; Pred. No. 7.51e+00;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 501 SSYLPTRKALD 511
:||||:|
QY 1 ASYLSTSSSLD 11

RESULT 9
ID 028628; PRELIMINARY; PRT; 1022 AA.
AC 028628;

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE A-KINASE ANCHORING PROTEIN AKAP120.
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUKARYOTA; METAZOA; CHORATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-NEW ZEALAND WHITE; TISSUE-STOMACH;
RC MEDLINE: 97220389.
RA DRANSFIELD D.T., YEH J.L., BRADFORD A.J., GOLDENRING J.R.;
RT "Identification and characterization of a novel A-kinase-anchoring
RT protein (AKAP120) from rabbit gastric parietal cells.";
RL BIOCHEM. J. 322:0-0(0).
DR EMBL: U26360; G35959584; -
SQ SEQUENCE 1022 AA; 116903 MW; 98BD43C2 CRC32;

Query Match 71.0%; Score 49; DB 6; Length 1022;
Best Local Similarity 72.7%; Pred. No. 7.51e+00;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 935 ASYLSTSSSLD 945
:||||:|
QY 2 SYLSTSSSLD 12

RESULT 10
ID 027302; PRELIMINARY; PRT; 160 AA.
AC 027302;

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE GLOBIN.
GN CBG1.
OS CAENORHABDITIS BRIGGSAE.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RA KLOER A.P., GOLDBERG D.E.;
RL SUBMITTED (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U48291; G1203913; -
DR EMBL: U48290; G1203913; JOINED.
DR EMBL: U48289; G1203909; -
DR PFAM: PF00042; globin; 1.
SQ SEQUENCE 160 AA; 18591 MW; 618F4F46 CRC32;

Query Match 69.6%; Score 48; DB 5; Length 160;

Best Local Similarity 41.7%; Pred. No. 1.24e+01;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
Db 119 TGYLESTGSLND 130
QY 1 ASYLSTSSSLDD 12

RESULT 11
ID Q00510 PRELIMINARY; PRT; 239 AA.
AC Q00510;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SRM PROTEIN.
GN SRM.
OS STREPTOMYCES AMBOFACIENS.
OC BACTERIA: FIRMICUTES: ACTINOBACTERIA: ACTINOBACTERIAE:
OC ACTINOMYCETALES: STREPTOMICINAE: STREPTOMYCETACEAE: STREPTOMYCES.
[1]
SEQUENCE FROM N.A.
STRAIN-BES2281;
RX MEDLINE: 92374852.
RA GESTLICH M., LOSICK R., TURNER J.R., RAU R.N.;
RT "Characterization of a novel regulatory gene governing the expression
of a polyketide synthase gene in Streptomyces ambofaciens";
RL MOL. MICROBIOL. 6:2019-2029(1992).
DR EMBL: X63451; G46702;
SQ SEQUENCE 239 AA; 26493 MW; 45A8C51A CRC32;

Query Match 69.6%; Score 48; DB 2; Length 239;
Best Local Similarity 45.5%; Pred. No. 1.24e+01;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Db 110 GYLETADLE 120
QY 2 SYLSTSSSLDD 12

RESULT 12
ID Q23173 PRELIMINARY; PRT; 484 AA.
AC Q23173;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE WOSELO.1 PROTEIN.
GN WOSELO.1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA: METAZOA: NEMATODA: SECERNENTEA: RHABDITIA: RHABDITIDA;
OC RHABDITINA: RHABDITOIDEA: RHABDITIDAE: PELODERINAE: CAENORHABDITIS.
[1]
SEQUENCE FROM N.A.
MORTIMORE B.;
RA SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
SEQUENCE FROM N.A.
MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BOFIELD J., BURTON J., CONNELL M., COSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans";
RL NATURE 368:32-38(1994).
DR EMBL: Z77670; E1350053;
SQ SEQUENCE 484 AA; 53896 MW; EC00CF7B CRC32;

Query Match 69.6%; Score 48; DB 5; Length 484;
Best Local Similarity 45.5%; Pred. No. 1.24e+01;
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
Db 150 AYVASSTLED 160
QY 2 SYLSTSSSLDD 12

RESULT 13
ID Q60167 PRELIMINARY; PRT; 649 AA.
AC Q60167;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE PROTEIN COMPLEX ASSEMBLY PROTEIN.
GN SPBC19F8.03C.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST)
OC EUKARYOTA: FUNGI: ASCOMYCOTA: ARCHIASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES: SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
[1]
SEQUENCE FROM N.A.
RC STRAIN-972H-;
RA BECK A., REINHARDT R., WOOD V., RAJANDREAM M.A., BARRELL B.G.;
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AL023594; E1293401;
SQ SEQUENCE 649 AA; 72985 MW; 9C207DB2 CRC32;

Query Match 69.6%; Score 48; DB 3; Length 649;
Best Local Similarity 60.0%; Pred. No. 1.24e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 234 NYLSTARSL 243
QY 2 SYLSTSSSLDD 11

RESULT 14
ID Q22830 PRELIMINARY; PRT; 1758 AA.
AC Q22830;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SIMILAR TO HUMAN SREBP-2 BASIC-HELIX-LOOP-HELIX-LEUCINE ZIPPER
DE TRANSCRIPTION FACTOR.
GN T27B1.1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA: METAZOA: NEMATODA: SECERNENTEA: RHABDITIA: RHABDITIDA;
OC RHABDITINA: RHABDITOIDEA: RHABDITIDAE: PELODERINAE: CAENORHABDITIS.
[1]
SEQUENCE FROM N.A.
MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BOFIELD J., BURTON J., CONNELL M., COSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans";
RL NATURE 368:32-38(1994).
[2]
SEQUENCE FROM N.A.
WU X., LE T.;
RA SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
[3]
SEQUENCE FROM N.A.
WATERSTON R.;

RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U41020; G1086678; -
 SQ SEQUENCE 1758 AA; 198406 MW; A58AC1E4 CRC32;

Query Match 69.6%; Score 48; DB 5; Length 1758;
 Best Local Similarity 41.7%; Pred. No. 1.24e+01;
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 962 SSYLAKEADLDE 973
 :|||: ||:
 QY 1 ASYLTSSSLDD 12

RESULT 15
 ID 068946 PRELIMINARY; PRT; 112 AA.
 AC 068946:
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DT DINITROGENASE 3 DELTA SUBUNIT.
 ANFG.

AZOMONAS MACROCYTOGENES.
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; AZOTOBACTERACEAE;
 OC AZOMONAS.

RN [1]
 RP SEQUENCE FROM N.A.
 RA LOVELESS T.M., BISHOP P.E.;
 RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF058780; G3065908; -
 SQ SEQUENCE 112 AA; 13273 MW; 5ABC839C CRC32;

Query Match 68.1%; Score 47; DB 2; Length 112;
 Best Local Similarity 50.0%; Pred. No. 2.03e+01;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 96 YLTSSSLDD 105
 ||: |||:
 QY 3 YLTSSSLDD 12

Search completed: Thu Sep 2 12:36:46 1999
 Job time : 23 secs.

THIS PAGE BLANK (USPTO)

WATERMAN

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch.p protein - protein database search, using Smith-Waterman algorithm

on: Thu Sep 2 12:35:58 1999; Maspar time 2.17 Seconds
156,010 Million cell updates/sec
Tabular output not generated.

Title: >US-08-599-226-30
Description: (1-12) from US08599226.pap
Sequence: 1 ASYLSTSSSLDD 12

Scoring table: PAM 150
Gap 15

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot37
I:swissprot

Statistics: Mean 24.293; Variance 24.652; scale 0.985

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	49	71.0	345	1	VCA_BPT7	MAJOR CAPSID PROTEIN 1
2	49	71.0	352	1	YDH3_SCHPO	HYPOTHETICAL 39.7 KD P
3	49	71.0	398	1	VCAB_BPT7	MINOR CAPSID PROTEIN 1
4	48	69.6	396	1	LDD_ECOLI	L-LACTATE DEHYDROGENAS
5	48	69.6	483	1	KICL_MOUSE	KERATIN, TYPE I CYTOSK
6	48	69.6	687	1	VITA_VIBCH	VIBRIOACETIN RECEPTOR
7	48	69.6	750	1	YJN2_YEAST	HYPOTHETICAL 84.5 KD P
8	47	68.1	213	1	GACD_PSEFL	RESPONSE REGULATOR GAC
9	47	68.1	333	1	YASD_MYCSM	HYPOTHETICAL 35.9 KD P
10	47	68.1	471	1	UF01_MAIZE	FLAVONOL 3-O-GLUCOSYL
11	47	68.1	471	1	UF03_MAIZE	FLAVONOL 3-O-GLUCOSYL
12	47	68.1	471	1	UF02_MAIZE	FLAVONOL 3-O-GLUCOSYL
13	47	68.1	578	1	HLI4_ABRSA	HEMOLYSIN 4 PRECURSOR
14	47	68.1	684	1	YV18_MYCTU	HYPOTHETICAL 57.3 KD P
15	46	66.7	115	1	NIFW_AZOVI	NIFW PROTEIN
16	46	66.7	127	1	MBP_RAT	MYELIN BASIC PROTEIN S
17	46	66.7	132	1	YV07_MYCTU	HYPOTHETICAL 14.7 KD P
18	46	66.7	167	1	MBP_CAVPO	MYELIN BASIC PROTEIN (
19	46	66.7	171	1	MBP_PANTR	MYELIN BASIC PROTEIN (
20	46	66.7	194	1	MBP_MOUSE	MYELIN BASIC PROTEIN (
21	46	66.7	196	1	MBP_HUMAN	MYELIN BASIC PROTEIN (
22	46	66.7	301	1	P34_RICRI	PROTEIN P34
23	46	66.7	347	1	VCA_BPT3	MAJOR CAPSID PROTEIN 1

24	46	66.7	411	1	KICL_RABIT	KERATIN, TYPE I CYTOSK	1.45e+01
25	46	66.7	433	1	VCAB_BPT3	MINOR CAPSID PROTEIN 1	1.45e+01
26	46	66.7	466	1	FLID_SALTY	FLINGULAR HOOK-ASSOCIA	1.45e+01
27	46	66.7	577	1	PYRH_YEAST	CTP SYNTHASE 2 (EC 6.3	1.45e+01
28	46	66.7	1839	1	ANKC_HUMAN	ANKRIN, BRAIN VARIANT	1.45e+01
29	46	66.7	3924	1	ANKB_HUMAN	ANKRIN, BRAIN VARIANT	1.45e+01
30	45	65.2	132	1	ANFG_AZOVI	NITROGENASE IRON-IRON	2.40e+01
31	45	65.2	204	1	UREG_STRL	UREASE ACCESSORY PROTE	2.40e+01
32	45	65.2	401	1	KICS_POTR	KERATIN, TYPE I CYTOSK	2.40e+01
33	45	65.2	416	1	YEIJ_ECOLI	HYPOTHETICAL 43.4 KD P	2.40e+01
34	45	65.2	431	1	KICQ_HUMAN	KERATIN, TYPE I CYTOSK	2.40e+01
35	45	65.2	456	1	KICQ_HUMAN	KERATIN, TYPE I CYTOSK	2.40e+01
36	45	65.2	469	1	KICX_HUMAN	KERATIN, TYPE I CYTOSK	2.40e+01
37	45	65.2	471	1	KICN_HUMAN	KERATIN, TYPE I CYTOSK	2.40e+01
38	45	65.2	494	1	KICL_HUMAN	KERATIN, TYPE I CYTOSK	2.40e+01
39	45	65.2	526	1	KICJ_BOVIN	KERATIN, TYPE I CYTOSK	2.40e+01
40	45	65.2	543	1	YVFB_YEAST	HYPOTHETICAL 60.8 KD P	2.40e+01
41	45	65.2	547	1	MERA_STAVU	MERCURIC REDUCTASE (EC	2.40e+01
42	45	65.2	593	1	KICJ_HUMAN	KERATIN, TYPE I CYTOSK	2.40e+01
43	45	65.2	633	1	YIJ2_YEAST	HYPOTHETICAL 71.0 KD P	2.40e+01
44	45	65.2	753	1	YBHJ_ECOLI	HYPOTHETICAL 81.5 KD P	2.40e+01
45	45	65.2	1298	1	YTFN_HAETN	HYPOTHETICAL PROTEIN H	2.40e+01

ALIGNMENTS

RESULT	ID	Score	Match	Length	ID	Description	Pred. No.
1	VCA_BPT7	71.0	345	1	VCA_BPT7	MAJOR CAPSID PROTEIN 1	1.45e+01
2	P19726; P03717;	71.0	352	1	YDH3_SCHPO	HYPOTHETICAL 39.7 KD P	1.45e+01
3	21-JUL-1986 (REL. 01, CREATED)	71.0	398	1	VCAB_BPT7	MINOR CAPSID PROTEIN 1	1.45e+01
4	01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)	71.0	396	1	LDD_ECOLI	L-LACTATE DEHYDROGENAS	1.45e+01
5	01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)	71.0	483	1	KICL_MOUSE	KERATIN, TYPE I CYTOSK	1.45e+01
6	MAJOR CAPSID PROTEIN 10A.	71.0	687	1	VITA_VIBCH	VIBRIOACETIN RECEPTOR	1.45e+01
7	10.	71.0	750	1	YJN2_YEAST	HYPOTHETICAL 84.5 KD P	1.45e+01
8	BACTERIOPHAGE T7.	71.0	213	1	GACD_PSEFL	RESPONSE REGULATOR GAC	1.45e+01
9	VIRUSES; DSDNA VIRUSES, NO RNA STAGE; TAILED PHAGES; PODOVIRIDAE.	71.0	333	1	YASD_MYCSM	HYPOTHETICAL 35.9 KD P	1.45e+01
10	[1]	71.0	471	1	UF01_MAIZE	FLAVONOL 3-O-GLUCOSYL	1.45e+01
11	SEQUENCE FROM N.A.	71.0	471	1	UF03_MAIZE	FLAVONOL 3-O-GLUCOSYL	1.45e+01
12	MEDLINE; 83241725.	71.0	471	1	UF02_MAIZE	FLAVONOL 3-O-GLUCOSYL	1.45e+01
13	DUNN J.J., STUDIER F.W.;	71.0	578	1	HLI4_ABRSA	HEMOLYSIN 4 PRECURSOR	1.45e+01
14	"Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements."	71.0	684	1	YV18_MYCTU	HYPOTHETICAL 57.3 KD P	1.45e+01
15	J. MOL. BIOL. 166:477-535(1983).	71.0	115	1	NIFW_AZOVI	NIFW PROTEIN	1.45e+01
16	-1- SIMILARITY: TO THE T3 MAJOR CAPSID PROTEIN 10A.	71.0	127	1	MBP_RAT	MYELIN BASIC PROTEIN S	1.45e+01
17	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	71.0	167	1	MBP_CAVPO	MYELIN BASIC PROTEIN (1.45e+01
18	EMBL; V01146; G15604; -	71.0	171	1	MBP_PANTR	MYELIN BASIC PROTEIN (1.45e+01
19	PIR; A04344; VABPA7.	71.0	194	1	MBP_MOUSE	MYELIN BASIC PROTEIN (1.45e+01
20	PIR; S42325; S42325.	71.0	196	1	MBP_HUMAN	MYELIN BASIC PROTEIN (1.45e+01
21	COAT PROTEIN.	71.0	301	1	P34_RICRI	PROTEIN P34	1.45e+01
22	SEQUENCE 345 AA; 36545 MW; D12C1BAE CRC32;	71.0	347	1	VCA_BPT3	MAJOR CAPSID PROTEIN 1	1.45e+01
23	Query Match	71.0					
24	Best Local Similarity	71.0					
25	Matches	6	Conservative	4	Mismatches	2	Indels
26							
27							
28							
29							
30							
31							
32							
33							
34							
35							
36							
37							
38							
39							
40							
41							
42							
43							
44							
45							

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 39.7 KD PROTEIN C6G9.03C IN CHROMOSOME I.
GN SPAC6G9.03C.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., CONNOR R.E.;
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/CCDB DATA BANKS.
CC -1- SIMILARITY: SOME, TO YEAST YNL206C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: 281317; E276610; -
KM HYPOTHETICAL PROTEIN.
SQ SEQUENCE 352 AA; 39679 MW; D92A9357 CRC32;

Query Match 71.0%; Score 49; DB 1; Length 352;
Best Local Similarity 54.5%; Pred. No. 3.05e+00;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 303 STYLNENASSEN 313
OY 2 STYLNENASSEN 12
ID VCAR_BPT7 STANDARD; PRT; 398 AA.
AC P19727; P03717;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE MINOR CAPSID PROTEIN 10B.
GN 10.
OS BACTERIOPHAGE T7.
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; TAILED PHAGES; PODOVIRIDAE.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 83241725.
RA DUNN J.J., THOMPSON K.;
RL "Complete nucleotide sequence of bacteriophage T7 DNA and the
RL locations of 77 genetic elements."
RT J. MOL. BIOL. 166:477-535(1983).
CC -1- THE MINOR CAPSID PROTEIN 10B IS DUE TO A TRANSLATIONAL SHIFT TO
CC THE -1 FRAME.
CC -1- SIMILARITY: TO THE T3 MINOR CAPSID PROTEIN 10B, EXCEPT IN THE
CC C-TERMINAL EXTENSIONS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: V01146; G431193; -
DR PIR: B04344; VBBA7.
DR PIR: S42326; S42326.
SQ COAT PROTEIN.
SQ SEQUENCE 398 AA; 41830 MW; 8BD410FF CRC32;

Query Match 71.0%; Score 49; DB 1; Length 398;
Best Local Similarity 50.0%; Pred. No. 3.05e+00;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 71 AAYLAGENLDD 82
OY 1 AAYLAGENLDD 12
ID LIDD_ECOLI STANDARD; PRT; 396 AA.
AC P33332;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE L-LACTATE DEHYDROGENASE (CYTOCHROME) (EC 1.1.2.3).
GN LIDD OR LCTD.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE: 94012541.
RA DONG J.M., TAYLOR J.S., LATOUR D.J., IUCHI S., LIN E.C.C.;
RT "Three overlapping lct genes involved in L-lactate utilization by
RT Escherichia coli."
RL J. BACTERIOL. 175:6671-6678(1993).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE: 94316500.
RA SORIA H.J., BURLAND V., DANIELS D.L., PLUNKETT G. III, BLATTNER F.R.;
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
RT region from 76.0 to 81.5 minutes."
RL NUCLEIC ACIDS RES. 22:2576-2586(1994).
CC -1- CATALYTIC ACTIVITY: (S)-LACTATE + 2 FERRICYTOCHROME C -> PYRUVATE +
CC 2 FERROCYTOCHROME C.
CC -1- COFACTOR: FMN.
CC -1- INDUCTION: AEROBICALLY BY L-LACTATE.
CC -1- SIMILARITY: BELONGS TO THE FMN-DEPENDENT ALPHA-HYDROXY ACID
CC DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L13970; G404695; -
DR EMBL: U00039; G466743; -
DR EMBL: AE000438; G1790033; -
DR PIR: C49904; C49904.
CC ECOCENE; B611963; LIDD.
DR PROSITE; P500557; FMN_HYDROXY_ACID_DH; 1.
DR PFAM; PF01070; FMN_ch; 1.
DR HSSP; P05414; IGYL.
KW OXIDOREDUCTASE; FLAVOPROTEIN; FMN.
FT ACT_SITE 24 SUBSTRATE BINDING (BY SIMILARITY).
FT ACT_SITE 129 SUBSTRATE BINDING (BY SIMILARITY).
FT ACT_SITE 275 REMOVES THE SUBSTRATE ALPHA-PROTON AS THE
FT FIRST STEP IN CATALYSIS (BY SIMILARITY).
FT ACT_SITE 278 SUBSTRATE BINDING (BY SIMILARITY).
SQ SEQUENCE 396 AA; 42728 MW; 10BAD3D1 CRC32;

Query Match 69.6%; Score 48; DB 1; Length 396;
Best Local Similarity 33.3%; Pred. No. 5.18e+00;
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Db 208 SAYLKGPTGLD 219
OY 1 SAYLKGPTGLD 12

RESULT 5
ID K1CL_MOUSE STANDARD: PRT: 483 AA.
AC 064291;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE KERATIN, TYPE I CYTOSKELETAL 12 (CYTOKERATIN 12).
GN KRT12 OR KRT1-12 OR KRT1.12.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
RN RODENTIA; SCIRURGICANTH; MORIDAE; MORINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=CORNEA;
RX MEDLINE: 94139368.
RA LIU C.-Y., ZHU G., WESTERHAUSEN-LARSON A., CONVERSE R., KAO C.W.-C.,
RA SUN T.-T., KAO W.W.-Y.;
RA "Cornea-specific expression of K12 keratin during mouse development.";
RT CURR. EYE RES. 12:963-974(1993).
[2]
SEQUENCE FROM N.A.
RC STRAIN=129/SVJ; TISSUE=CORNEA;
RX MEDLINE: 95014223.
RA LIU C.-Y., ZHU G., CONVERSE R., KAO C.W.-C., NAKAMURA H.,
RA TSENG S.C.-G., MUI M.-M., SEYER J., JUSTICE M.J., STECH M.E.,
RA HANSEN G.M., KAO W.W.-Y.;
RA "Characterization and chromosomal localization of the cornea-specific
murine keratin gene Krt12.12.";
RT J. BIOL. CHEM. 269:24627-24636(1994).
RL [1]
CC - FUNCTION: MAY PLAY A UNIQUE ROLE IN MAINTAINING THE NORMAL CORNEAL
EPITHELIAL FUNCTION.
CC - SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC - KERATIN 3 ASSOCIATES WITH KERATIN 12.
CC - TISSUE SPECIFICITY: CORNEA-SPECIFIC. ASSOCIATED MAINLY WITH ALL
LAYERS OF THE CENTRAL CORNEAL EPITHELIUM AND ALSO FOUND IN THE
SUPRABASAL LIMBAL EPITHELIUM.
CC - THERE ARE TWO TYPES OF CYTOSKELETAL AND MICROFIBRILLAR KERATIN,
I (ACIDIC) AND II (NEUTRAL TO BASIC) (40-55 AND 56-70 KILODALTONS,
RESPECTIVELY).
CC - ALTERNATIVE PRODUCTS: DIFFERENT FORMS MAY BE PRODUCED BY
ALTERNATIVE SPLICING.
CC - SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: U02880; G414540; -;
CC EMBL: U08095; G565660; -;
CC MGI: 96687; KRT1-12.
CC PROSITE: PS00226; IF: 1.
CC PRAM: PF00038; filament; 1.
CC KM INTERMEDIATE FILAMENT; COILED COIL; HEPTAD REPEAT PATTERN; KERATIN;
KW ALTERNATIVE SPLICING.
CC
CC FT DOMAIN 1 118 HEAD.
CC FT DOMAIN 119 428 ROD.
CC FT DOMAIN 429 483 TAIL.
CC FT DOMAIN 119 154 COIL 1A.
CC FT DOMAIN 158 175 LINKER 1.
CC FT DOMAIN 176 267 COIL 1B.
CC FT DOMAIN 268 290 LINKER 12.
CC FT DOMAIN 291 428 COIL 2.
CC FT VARSPLIC 211 212 YE -> KL (IN K12-ALT9).
CC VARSPLIC 213 483 MISSING (IN K12-ALT9).
CC SO SEQUENCE 483 AA; 52010 MW; 54DB7635 CRC32;
Query Match 69.6%; Score 48; DB 1; Length 483;
Best Local Similarity 50.0%; Pred. No. 5.18e+00;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Db 131 ASYLKVRSLEE 142
|||:|:|:
QY 1 ASYLSTSSLD 12
RESULT 6
ID VIUA_VIBCH STANDARD: PRT: 687 AA.
AC 000964;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE VIBRIOACTIN RECEPTOR PRECURSOR.
GN VIUA.
OS VIBRIO CHOLERAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; VIBRIONACEAE; VIBRIO.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OGAMA 395;
RX MEDLINE: 92276356.
RA BUTTERTON J.R., STOEBERNER J.A., PAYNE S.M., CALDERWOOD S.B.;
RT "Cloning, sequencing, and transcriptional regulation of viua, the
gene encoding the ferric vibriobactin receptor of Vibrio cholerae.";
RL J. BACTERIOL. 174:3729-3738(1992).
[2]
SEQUENCE FROM N.A.
RC STRAIN=OGAMA 395;
RA LIAO W.J., CHOI M.H., BUTTERTON J.R.;
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC - FUNCTION: RECEPTOR FOR FERRIC VIBRIOBACTIN.
CC - SUBCELLULAR LOCATION: OUTER MEMBRANE.
CC -
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: AF030977; G2641147; -;
CC DR PIR: A41905; A41905.
CC PRAM: PF00593; TonB box; 1.
CC KM SIGNAL; RECEPTOR; OUTER MEMBRANE; IRON TRANSPORT.
CC FT SIGNAL 1 37
CC FT CHAIN 38 687 VIBRIOBACTIN RECEPTOR.
CC SO SEQUENCE 687 AA; 76413 MW; 525D3D49 CRC32;
Query Match 69.6%; Score 48; DB 1; Length 687;
Best Local Similarity 60.0%; Pred. No. 5.18e+00;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Db 662 YLSTNTLDQ 671
|||:|:|:
QY 3 YLSTSSLD 12
RESULT 7
ID YUN2_YEAST STANDARD: PRT: 750 AA.
AC P47014;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 84.5 KD PROTEIN IN MR3-URA2 INTERGENIC REGION.
GN YJ132W OR J0678.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
RN SACCCHAROMYCETACEAE; SACCCHAROMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE: 96408771.

```

RA KANSOULOU C., TZEREMIA M., TAVENAKAKIS N., ALEXANDRAKI D.;
RT "Sequence analysis of a 40.7 kb segment from the left arm of yeast
RT chromosome X reveals 14 known genes and 13 new open reading frames
RT including homologues of genes clustered on the right arm of
RT chromosome XI.";
RL YEAST 12:787-797(1996).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: SOME: NO MAMMALIAN PHOSPHATIDYLINOSITOL-GLYCAN-
CC SPECIFIC PHOSPHOLIPASE D (EC 3.1.4.50).
CC -----
CC CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/).
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X87371: G854564: -.
DR EMBL: Z49407: G1008336: -.
DR HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
RT TRANSMEM 1 21 POTENTIAL.
RT TRANSMEM 465 485 POTENTIAL.
RT TRANSMEM 586 606 POTENTIAL.
SO SEQUENCE 750 AA: 84466 MW: 5D6D701 CRC32;

Query Match 69.6%; Score 48; DB 1; Length 750;
Best Local Similarity 50.0%; Pred. No. 5.18e+00;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 354 YLAVSAPLED 363
YL 1:1:1:1:1
YL 3 YLSTSSLD 12

RESULT 8 STANDARD: PRT: 213 AA.
ID GACA_PSEFL GACA_PSEFL
AC G32987;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE RESPONSE REGULATOR GACA (GLOBAL ACTIVATOR) (GLOBAL ANTIBIOTIC AND
DE CYANIDE CONTROL PROTEIN).
GN GACA.
OS PSEUDOMONAS FLUORESCENS.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;
OC PSEUDOMONAS.
RN [1]
RT SEQUENCE FROM N.A.
RT STRAIN=CHAO;
RX MEDLINE: 92179223.
RA LAVILLE J., VOISARD C.P., KEEL C., MAURHOEFER M., DIFAGO G.,
RA HAAS D.;
RT "Global control in pseudomonas fluorescens mediating antibiotic
RT synthesis and suppression of black root rot of tobacco.";
RL PROC. NATL. ACAD. SCI. U.S.A. 89:1562-1566(1992).
RL [2]
RN SEQUENCE FROM N.A.
RN STRAIN=BL915;
RP MEDLINE: 94355677.
RX GAFFNEY T.D., LAM S.T., LISON J., GATES K., FRAZELLE A., MAIO J.,
RA HILL S., GOODWIN S., TORKEWITZ N., ALLSHOUSE A.M., KEMPF H.J.,
RA BECKER J.O.;
RT "Global regulation of expression of antifungal factors by a
RT pseudomonas fluorescens biological control strain.";
RL MOL. PLANT MICROBE INTERACT. 7:455-463(1994).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF SECONDARY METABOLISM.
CC INVOLVED IN THE SYNTHESIS OF THE ANTIFUNGAL FACTORS CYANIDE AND
CC 2,4-DIACEYLDIHYDROGLUCINOL.
CC -1- SIMILARITY: TO OTHER BACTERIAL REGULATORY PROTEINS INVOLVED IN
CC SIGNAL TRANSDUCTION
CC -1- SIMILARITY: BELONGS TO THE LUXR/DHPA FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.

```

```
CC ----- This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its use  
CC by non-profit institutions as long as its content is in no way used  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; M80913; G551930; -.  
DR EMBL; L29642; G472403; -.  
DR PROSITE; PS00622; HTH_LUXR_FAMILY; 1.  
DR PFAM; PF00196; GERE; 1.  
DR HSSP; P10957; INRL.  
KW TRANSCRIPTION REGULATION; DNA-BINDING; SENSORY TRANSDUCTION;  
KW PHOSPHORYLATION.  
FT DOMAIN 1 99 RECEPTOR DOMAIN.  
FT MOD_RES 54 54 PHOSPHORYLATION (BY SIMILARITY).  
FT DNAS_BIND 166 185 H-T-H MOTIF (POTENTIAL).  
FT VARIANT 90 90 P -> L (IN THE SPONTANEOUS PLEIOTROPIC  
FT MUTANT BL915-1).  
FT VARIANT 182 182 T -> I (IN THE SPONTANEOUS PLEIOTROPIC  
FT MUTANT BL915-2).  
FT CONFLICT 49 49 Y -> D (IN REF. 2).  
FT CONFLICT 213 213 L -> A (IN REF. 2).  
SQ SEQUENCE 213 AA; 23454 MW; 42DC9AC2 CRC32;  
  
Query Match 68.1%; Score 47; DB 1; Length 213;  
Best Local Similarity 33.3%; Pred. No. 8.71e+00;  
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;  
  
Db 99 AGYLRKAGALNE 110  
|::|:::::  
Qy 1 ASYSTSSLD 12  
  
RESULT 9  
ID YASD_MYCSM STANDARD; PTR; 333 AA.  
AC P41402;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DE 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE HYPOTHEICAL 35.9 KD PROTEIN IN ASD 3'REGION (ORFY).  
OS MYCOBACTERIUM SMEGMATIS.  
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIIDE;  
OC ACINOMYCETALES; CORYNBACTERINEAE; MYCOBACTERACEAE; MYCOBACTERIUM.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 607 / MC(2)6;  
RX MEDLINE; 94254720.  
RA CIRILLO J.D.; WEISBROD T.R.; PASOPELLA L.; BLOOM B.R.,  
RA JACOBS W.R. JR.;  
RT "Isolation and characterization of the aspartokinase and aspartate  
RT semialdehyde dehydrogenase operon from mycobacteria.";  
RL MOL. MICROBIOL. 11:629-639(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; Z17372; G581353; -.  
KW HYPOTHEICAL PROTEIN.  
SQ SEQUENCE 333 AA; 35881 MW; 52FCSB23 CRC32;  
  
Query Match 68.1%; Score 47; DB 1; Length 333;  
Best Local Similarity 41.7%; Pred. No. 8.71e+00;  
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
```


CC GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
CC PIGMENTS.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + A FLAVONOL = UDP + FLAVONOL
CC 3-O-D-GLUCOSIDE.
CC -1- PATHWAY: ONE OF THE ESSENTIAL AND TERMINAL STEPS IN THE
CC ANTHOCYANIN BIOSYNTHETIC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X13501; G295854; -
CC PIR: S08325; S08325.
CC MAI:EDB: 13885; -
CC PROSITE: PS00375; UDEPT: 1.
CC PFM: PFO0201; UDEPT: 2.
CC TRANSFERASE: GLYCOSYLTRANSFERASE.
CC SEQUENCE 471 AA: 48621 MW: 3158C5E0 CRC32;

Query Match 68.1%; Score 47; DB 1; Length 471;
Best Local Similarity 63.6%; Pred. No. 8.71e+00;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 47 SFLSTRASSIAQ 57
1:1111111111
Qy 2 SYLSTSSSLDD 12

RESULT 13
ID HLX4_AERSA STANDARD; PRT: 578 AA.
AC 008677;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HEMOLYSIN 4 PRECURSOR.
GN ASH4.
OS AEROMONAS SALMONICIDA.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; AEROMONAS GROUP;
OC AEROMONAS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-17-2;
RX MEDLINE: 94142497.
RT HIRONO I., AOKI T.;
RT "Cloning and characterization of three hemolysin genes from Aeromonas
RT salmonicida."
RT MICROB. PATHOG. 15:269-282(1993).
CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC DEFINED.
CC -1- SIMILARITY: BELONGS TO THE AHN1/ASH4/HLXA/VVHA FAMILY OF
CC HEMOLYSINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X65049; G39012; -
CC PFM: PFO0652; R1cin_B_lectin; 1.
CC HEMOLYSIN: TOXIN; SIGNAL.
CC SIGNAL 1 578 POTENTIAL.
CC CHAIN ? 578 HEMOLYSIN 4.
CC SEQUENCE 578 AA: 63400 MW: 8A23C8BC CRC32;

Query Match 68.1%; Score 47; DB 1; Length 578;
Best Local Similarity 60.0%; Pred. No. 8.71e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 13 YINASSWIDE 22
1:1111111111
Qy 3 YLSTSSSLDD 12

RESULT 14
ID YV18_MYCTU STANDARD; PRT: 684 AA.
AC Q1157;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 57.3 KD PROTEIN GMC-TYPE OXIDOREDUCTASE CY20G9.18C.
GN MYC20G9.18C.
OS MYCOBACTERIUM TUBERCULOSIS.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
OC ACTINOMYCETALES; CORYNEBACTERIINEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-H37RV;
RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- COFACTOR: FAD FLAVOPROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z77162; E255032; -
CC DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 1.
CC DR PROSITE: PS00624; GMC_OXRED.2; 1.
CC KW HYPOTHETICAL PROTEIN; OXIDOREDUCTASE; FLAVOPROTEIN; FAD.
CC SEQUENCE 684 AA: 72253 MW: F4A66D1F CRC32;

Query Match 68.1%; Score 47; DB 1; Length 684;
Best Local Similarity 70.0%; Pred. No. 8.71e+00;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 117 ASYLTTGRSL 126
1111111111
Qy 1 ASYLSTSSSL 10

RESULT 15
ID N1FW_AZOVI STANDARD; PRT: 115 AA.
AC P1488;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE N1FW PROTEIN.
GN N1FW.
OS AZOTOBACTER VINELANDII.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; AZOTOBACTERACEAE;
OC AZOTOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 89123097.
RA JACOBSON M.R., BRIGLE K.E., BENNETT L.T., SETTERQUIST R.A.,
RA WILSON M.S., CASH V.L., BEYRON J., NEWTON W.E., DEAN D.R.;
RT "Physical and genetic map of the major nif gene cluster from
RT Azotobacter vinelandii."
RT J. BACTERIOL. 171:1017-1027(1989).
CC -1- SIMILARITY: BELONGS TO THE N1FW FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; M20568; G142365; -
CC PIR; D32055; D32055.
CC NITROGEN FIXATION.
CC KW
CC SEQUENCE 115 AA; 13418 MW; 74E9995C CRC32;

Query Match 66.7%; Score 46; DB 1; Length 115;
Best Local Similarity 50.0%; Pred. No. 1.45e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 54 YLSKAGDIDE 63
|||::||:
QY 3 YLSTSSSLDD 12

Search completed: Thu Sep 2 12:36:06 1999
Job time : 8 secs.

THIS PAGE BLANK (USPTO)

CC scleriosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA:

Query Match 100.0%; Score 82; DB 27; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.71e-01;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 asy1stsfslidy 12
| | | | | | | | | | | | | |
Qy 1 ASYSTSFSLDY 12

RESULT 2
ID W27588 standard; peptide: 12 AA.
AC W27588:
DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN WO9729131-A1.
PD 14-AUG-1997.
PE 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226.
PA (BANDI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Markovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity; e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 73; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA:

Query Match 81.7%; Score 67; DB 27; Length 12;
Best Local Similarity 91.7%; Pred. No. 8.27e+00;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 asy1stsfslidy 12
| | | | | | | | | | | | | |
Qy 1 ASYSTSFSLDY 12

RESULT 3
ID W27593 standard; peptide: 12 AA.
AC W27593:
DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN WO9729131-A1.
PD 14-AUG-1997.
PE 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226.
PA (BANDI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Markovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity; e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 75; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA:

Query Match 78.0%; Score 64; DB 27; Length 12;
Best Local Similarity 75.0%; Pred. No. 1.61e+01;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 1 asy1stsfslidy 12
| | | | | | | | | | | | | |
Qy 1 ASYSTSFSLDY 12

RESULT 4
ID W27591 standard; peptide: 12 AA.
AC W27591:
DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN WO9729131-A1.
PD 14-AUG-1997.
PE 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226.
PA (BANDI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Markovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity; e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 75; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA:

Query Match 78.0%; Score 64; DB 27; Length 12;
Best Local Similarity 75.0%; Pred. No. 1.61e+01;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

KM heavy chain: complementarity determining region 3; inhibition;
 KM treatment: sepsis; disease: autoimmune disease; infectious disease;
 KM malignancy: pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder: inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease: coagulation disturbance; burn; ELAM-1;
 KM keloid formation: scar tissue formation; pyrexia; HUVEC;
 KM periodontal disease: obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN W09729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997: U02219.
 PR 25-NOV-1996: US-031476.
 PI 09-FEB-1996: US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
 WPI: 97-415302/38.
 PS High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 Claim 20: Page 74, 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA;
 SQ

Query Match 74.4%; Score 61; DB 27; Length 12;
 Best Local Similarity 83.3%; Pred. No. 3.10e+01;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 asy1stss1ny 12
 ||||| | | |
 1 ASYSTSTSLDY 12

RESULT 5
 ID W27592 standard; peptide; 12 AA.
 AC W27592;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KM Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KM heavy chain: complementarity determining region 3; inhibition;
 KM treatment: sepsis; disease: autoimmune disease; infectious disease;
 KM malignancy: pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder: inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease: coagulation disturbance; burn; ELAM-1;
 KM keloid formation: scar tissue formation; pyrexia;
 KM periodontal disease: obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN W09729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997: U02219.

PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
 WPI: 97-415302/38.
 PS High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 Claim 20: Page 74, 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA;
 SQ

Query Match 74.4%; Score 61; DB 27; Length 12;
 Best Local Similarity 75.0%; Pred. No. 3.10e+01;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 as1stss1ey 12
 ||:|||| | | |
 1 ASYSTSTSLDY 12

RESULT 6
 ID W27569 standard; Protein; 121 AA.
 AC W27569;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain variable region.
 KM Human: tumour necrosis factor-alpha; TNF-alpha; antibody;
 KM heavy chain: variable region; inhibition;
 KM treatment: sepsis; disease: autoimmune disease; infectious disease;
 KM malignancy: pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder: inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease: coagulation disturbance; burn; ELAM-1;
 KM keloid formation: scar tissue formation; pyrexia; HUVEC;
 KM periodontal disease: obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN W09729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997: U02219.
 PR 25-NOV-1996: US-031476.
 PI 09-FEB-1996: US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TV, White M, Wilton AJ;
 WPI: 97-415302/38.
 PS High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 Claim 16: Page 76; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 factor-alpha (TNF-alpha) antibody (Ab) heavy chain variable region.

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 121 AA;

Query Match 74.4%; Score 61; DB 27; Length 121;
Best Local Similarity 81.8%; Pred. No. 3.10e+01;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 100 sylstssld 110
||| |||
Oy 2 sylstsfld 12

RESULT 7
ID W27586 standard; peptide: 12 AA.

AC W27586;

DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain CDR3.

KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.

OS Homo sapiens.

PN WO9729131-A1.

PD 14-AUG-1997.

PF 10-FEB-1997; U02219.

PR 25-NOV-1996; US-031476.

PT 09-FEB-1996; US-599226.

PI (BADI) BASF AG.

PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,
PI Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;

PI WPI: 97-41302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer

PS Claim 20; Page 72; 102pp; English.

CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity

CC determining region 3 (CDR3).

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or

CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro

CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,

CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple

CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,

CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,

CC burns, reperfusion injury, keloid formation, scar tissue formation,

CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

Query Match 67.1%; Score 55; DB 27; Length 12;
Best Local Similarity 90.9%; Pred. No. 1.12e+02;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 asylstssld 11
||| |||
Oy 1 asylstsfld 11

RESULT 8
ID W27589 standard; peptide: 12 AA.

AC W27589;

DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain CDR3.

KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.

OS Homo sapiens.

PN WO9729131-A1.

PD 14-AUG-1997.

PF 10-FEB-1997; U02219.

PR 25-NOV-1996; US-031476.

PT 09-FEB-1996; US-599226.

PI (BADI) BASF AG.

PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,
PI Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;

PI WPI: 97-41302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer

PS Claim 20; Page 73; 102pp; English.

CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity

CC determining region 3 (CDR3).

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or

CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro

CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,

CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple

CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,

CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,

CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The

CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein

CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

Query Match 67.1%; Score 55; DB 27; Length 12;
Best Local Similarity 90.9%; Pred. No. 1.12e+02;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 asylstssld 11
||| |||
Oy 1 asylstsfld 11

RESULT 9
ID W27587 standard: peptide: 12 AA.
AC W27587:
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW peridontal disease; obesity; radiation; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
14-AUG-1997.
10-FEB-1997: U02219.
25-NOV-1996: US-031476.
09-FEB-1996: US-599226.
PR (BAD1) BASF AG.
PA Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankevich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity; e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 73; 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption diseases,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, peridontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA:
Query Match 67.1%; Score 55; DB 27; Length 12;
Best Local Similarity 90.9%; Pred. No. 1.12e+02;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 1 asyistssld 11
QY 1 ASYLSTSFSLD 11

FR Key Location/Qualifiers
FT peptide 1..19
FT /note- "Signal peptide"
FT protein 20..141
FT /note- "Mature protein"
PN W09402627-A.
PD 03-FEB-1994.
PE 08-JUL-1993: J00941.
PR 16-JUL-1992: JP-189248.
PA (SAKA) OTSUKA PHARM CO LTD.
PI Hirai Y, Nishida T, Omoto Y, Owens RJ;
DR WPI: 94-048885/06.
N-PSDB: Q56066.
PT Mouse/human chimeric antibody against human interleukin-1 - for
PT treatment of diseases in which production of interleukin-1 is
PT abnormal, and for diagnostic imaging of interleukin-1 production
PT sites in vivo
PS Claim 1; Page 30-31; 58pp: Japanese.
CC The sequences given in R47205-08 represent the light and heavy chain,
CC variable and constant regions of a chimeric recombinant antibody
CC against human interleukin-1 (IL-1). The antibody has a light (L)
CC chain in which the constant region is that of a human antibody and
CC the variable region is from a mouse anti-human IL-1 antibody or is a
CC mouse-human graft containing the CDR regions of mouse anti-human IL-1
CC antibody, and a heavy (H) chain in which the constant region is that
CC of a human antibody and the variable region is from a mouse anti-human
CC IL-1 antibody or is a mouse-human graft containing the CDR regions of
CC mouse anti-human IL-1 antibody. The chimeric antibody is used to
CC treat diseases in which abnormal amounts of IL-1 are produced, eg.
CC inflammatory disease, arteriosclerosis, diffused intravascular
CC coagulation or leukemia. It can also be labelled and used for
CC diagnostic imaging of IL-1 producing sites in vivo.
SQ Sequence 141 AA:
Query Match 67.1%; Score 55; DB 9; Length 141;
Best Local Similarity 50.0%; Pred. No. 1.12e+02;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Db 121 yfgssyaldy 130
QY 3 YLSTSFSLDY 12

RESULT 11
ID R47208 standard: Protein: 147 AA.
AC R47208:
DT 09-AUG-1994 (first entry)
DE Human/murine IL-1 chimeric antibody VH.
KW Probe; chimeric; recombinant; antibody; human; interleukin-1; IL-1;
KW light; L: chain; constant; region; variable; mouse; anti-human;
KW graft; CDR; complementarity determining region; heavy; H;
KW inflammatory disease; arteriosclerosis; detection;
KW diffused intravascular coagulation; leukemia.
OS Synthetic.
FT Key Location/Qualifiers
FT peptide 1..19
FT /note- "Signal peptide"
FT protein 20..147
FT /note- "Mature protein"
PN W09402627-A.
PD 03-FEB-1994.
PE 08-JUL-1993: J00941.
PR 16-JUL-1992: JP-189248.
PA (SAKA) OTSUKA PHARM CO LTD.
PI Hirai Y, Nishida T, Omoto Y, Owens RJ;
DR N-PSDB: Q56069.
PT Mouse/human chimeric antibody against human interleukin-1 - for
PT treatment of diseases in which production of interleukin-1 is
PT abnormal, and for diagnostic imaging of interleukin-1 production
PT sites in vivo
PS Claim 2; Fig 13; 58pp: Japanese.
CC The sequences given in R47205-08 represent the light and heavy chain,

FT Misc_difference /note= "encoded by AAT"
 FT Misc_difference 659 /note= "encoded by TGC"
 FT Misc_difference 733 /note= "encoded by TGC"
 PN W09834950-A1.
 PD 13-AUG-1998.
 PF 06-FEB-1998; U02332.
 PR 07-FEB-1997; US-037859.
 PA (UYTE-) UNIV TENNESSEE RES CORP.
 PI Becker JM, Lubkowitz MA;
 DR WPI: 98-447166/38.
 DR N-PSDB: V49601.
 PT New oligopeptide membrane transporter from *Candida albicans* and
 PT related nucleic acid - used to deliver therapeutic agents,
 PT especially antifungals, to target cells and to identify plant
 PT transporters
 PS Claim 6; Fig 3a-c; 44p; English.
 CC This is the *Candida albicans* oligopeptide transporter, OPT1. Its
 CC amino acid sequence was deduced from an isolated genomic clone (see
 CC V49601) obtained through heterologous expression in *Saccharomyces*
 CC *cerevisiae* di-/tripeptide transport mutant PBIX-98. When
 CC transformed with a plasmid harboring OPT1, mutant PBIX-98, which
 CC does not express tetra-/pentapeptide transport activity under the
 CC conditions used, was conferred with an oligopeptide transport
 CC phenotype as indicated by growth on Lys-Leu-Gly, sensitivity to
 CC toxic tetra- and pentapeptides, and an increase in the initial
 CC uptake rate of the radiolabeled tetrapeptide Lys-Leu-Gly-(3H)Leu.
 CC The deduced protein product contains 12 hydrophobic regions,
 CC suggestive of a membrane transport protein. The oligopeptide
 CC transporter facilitates the uptake of normally non-permeable drugs,
 CC and is especially used to target antifungals, specifically
 CC anti-*Candida* drugs, which are conjugated, linked or attached to the
 CC peptide being transported. It can also be used to identify plant
 CC oligopeptide transporter genes, making it possible to use
 CC oligopeptides as herbicides or growth stimulants.
 SQ Sequence 783 AA;

Query Match 64.6%; Score 53; DB 35; Length 783;
 Best Local Similarity 50.0%; Pred. No. 1.71e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 406 flstfaisy 415
 :||:|:|:
 QY 3 YLSTSFSLDY 12

RESULT 15
 R83040 standard; peptide; 306 AA.
 R83040:
 DT 03-APR-1996 (first entry)
 DE Capsular polysaccharide of *Streptococcus pneumoniae* (CpsV).
 KW Capsular polysaccharide; cps; peptide; flanking region; detection;
 KW serotype; diagnosis; prevention; *Streptococcus pneumoniae*.
 OS *Streptococcus pneumoniae*.
 PN W09531548-A1.
 PD 23-NOV-1995.
 PF 16-MAY-1995; U06119.
 PR 16-MAY-1994; US-243546.
 PA (UABR-) UAB RES FOUND.
 PI Dillard JP, Yother J;
 DR WPI: 96-010934/01.
 DR N-PSDB: T05848.
 PT New *Streptococcus pneumoniae* capsular polysaccharide genes - used
 PT for detection, serotyping and for diagnosis and prevention of *S.*
 PT *pneumoniae* infection
 PS Disclosure: Page 178-180; 225pp; English.
 CC Sequences encoding the 5' flanking region of the capsular
 CC polysaccharide gene (cps) of *Streptococcus pneumoniae* and which are
 CC of sufficient length to allow hybridisation under standard
 CC hybridisation conditions to a *S.pneumoniae* cps gene flanking region
 CC may be used in methods to detect and serotype *S.pneumoniae*. They
 CC may also be used for the diagnosis and prevention of *S. pneumoniae*

CC infection. 306 AA;
 SQ Sequence
 Query Match 63.4%; Score 52; DB 15; Length 306;
 Best Local Similarity 40.0%; Pred. No. 2.11e+02;
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 67 yfstfley 76
 :||:|:|:
 QY 3 YLSTSFSLDY 12

Search completed: Thu Sep 2 12:38:34 1999
 Job time : 20 secs.

THIS PAGE BLANK (USPTO)

 M E S S A G E

 (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSPch_pp protein - protein database search, using Smith-Waterman algorithm

on: Thu Sep 2 12:39:53 1999; Maspar time 4.71 Seconds
 138.937 Million cell updates/sec

 Molecular output not generated.

Title: >US-08-599-226-31
 Description: (1-12) from US08599226.pep
 Perfect Score: 82
 Sequence: 1 ASYLSTSFSLDY 12

Scoring table: PAM 150
 Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

sptrembl9
 1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human
 5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.organelle
 9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified
 13:sp.vertebrate 14:sp.virus

Statistics: Mean 25.125; Variance 38.161; scale 0.658

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	63	76.8	372	5	P91143	6.49e+01
2	61	74.4	444	5	P91141	1.49e+00
3	59	72.0	472	2	086563	3.39e+00
4	59	72.0	1367	5	020120	3.39e+00
5	56	68.3	315	2	055611	1.12e+01
6	55	67.1	306	6	077768	1.65e+01
7	55	67.1	357	2	005237	1.65e+01
8	55	67.1	360	8	047575	1.65e+01
9	55	67.1	603	8	079437	1.65e+01
10	54	65.9	84	3	004203	2.43e+01
11	54	65.9	193	3	012064	2.43e+01
12	54	65.9	262	2	026039	2.43e+01
13	54	65.9	316	2	065998	2.43e+01
14	53	64.6	389	5	020272	3.56e+01
15	53	64.6	783	3	014411	3.56e+01
16	52	63.4	248	1	059270	5.18e+01
17	52	63.4	253	2	096151	5.18e+01
18	52	63.4	265	10	024226	5.18e+01
19	52	63.4	265	10	064442	5.18e+01
20	52	63.4	336	2	047450	5.18e+01

RESULT	ID	PRELIMINARY;	PRT;	372 AA.	ALIGNMENTS
21	52	63.4	336	2	047445
22	52	63.4	336	2	047444
23	52	63.4	337	2	047448
24	52	63.4	405	5	017660
25	52	63.4	450	2	024913
26	52	63.4	458	2	073215
27	52	63.4	551	11	088446
28	52	63.4	556	4	075751
29	52	63.4	588	2	096828
30	52	63.4	595	10	046675
31	52	63.4	649	2	084307
32	52	63.4	722	10	004514
33	51	62.2	118	10	081621
34	51	62.2	153	5	045791
35	51	62.2	189	10	022237
36	51	62.2	244	2	007432
37	51	62.2	341	5	018049
38	51	62.2	422	4	075462
39	51	62.2	428	5	017394
40	51	62.2	454	2	051495
41	51	62.2	582	5	022274
42	51	62.2	2014	5	001703
43	51	62.2	3391	14	089313
44	51	62.2	3391	14	089314
45	51	62.2	3391	14	089315

RESULT	ID	PRELIMINARY;	PRT;	372 AA.	ALIGNMENTS
1	P91143				
AC	P91143				
DT	01-MAY-1997 (TREMBL:REL. 03, CREATED)				
DT	01-MAY-1997 (TREMBL:REL. 03, LAST SEQUENCE UPDATE)				
DT	01-NOV-1998 (TREMBL:REL. 08, LAST ANNOTATION UPDATE)				
DE	SIMILAR TO ACETYLTRANSFERASES.				
GN	C37H5.2.				
OS	CAENORHABDITIS ELEGANS.				
OC	EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIDA; RHABDITIDA;				
OC	RHABDITIDA; RHABDITIDAE; RHABDITIDAE; PELDORINAE; CAENORHABDITIS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BRISTOL N2;				
RX	MEDLINE: 94150718.				
RA	WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,				
RA	BONFIELD J., BORTON J., CONNELL M., COSEY T., COOPER J., COULSON A.,				
RA	CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,				
RA	GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIRI M., JOHNSTON L.,				
RA	JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,				
RA	LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,				
RA	PARSONS J., PERCY C., KIRKEN L., ROOPRA A., SHUNDERS D., SHOWNKEEN R.,				
RA	SHALDON N., SMITH A., SONNHAMER E., STADEN R., SULTON J.,				
RA	THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,				
RA	WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLIDAN P.,				
RT	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.				
RT	elegans";				
RL	NATURE 368:32-38(1994).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BRISTOL N2;				
RA	DAVIDSON S., GILLAM B.,				
RL	SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.				
DR	EMBL: U88315; G1825777; -;				
DR	PRAM: P800561; abhydrolase; 1.				
KW	TRANSFERASE.				
SQ	SEQUENCE 372 AA; 42139 MW; 5214F159 CRC32;				

Query Match

76.8%; Score 63; DB 5; Length 372;

Best Local Similarity 58.3%; Pred. No. 6,49e-01;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 168 GYLSTSYALKY 179
Oy 1 ASYLSTFSLDY 12

RESULT 2
ID P91141 PRELIMINARY; PRT; 444 AA.

AC P91141;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SIMILAR TO ACETYLTRANSFERASES.
GN C37H5.3.

OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
RC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.

[1]
SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718.

RA WILSON R., AINSKOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,
RA BOURFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOMKNEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULLIVAN J.,
RA THERIAR-MEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans."
RL NATURE 368:32-38(1994).

[2]
SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;
RA DAVIDSON S., GILLAM B.;

RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;
RA WATERSTON R.;

RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U88315; G1825778; -.

DR PFAM: PF00561; abhydrolase_1.
RC TRANSFERASE.

SEQUENCE 444 AA; 50487 MW; F706E46B CRC32;

Query Match 74.4%; Score 61; DB 5; Length 444;
Best Local Similarity 50.0%; Pred. No. 1.49e+00;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 240 GYLATSYALKY 251
Oy 1 ASYLSTFSLDY 12

RESULT 3
ID 086563 PRELIMINARY; PRT; 472 AA.

AC 086563;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TRANSMEMBRANE TRANSPORT PROTEIN.
GN SC2A11.02C.

OS STREPTOMYCES COELICOLOR.

OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;

OC ACTINOMYCETALES; STREPTOMYCINAE; STREPTOMYCETACEAE; STREPTOMYCES.

RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);
RA MURPHY L., HARRIS D.;

RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;

RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.

RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RX MEDLINE: 97000351.

RA REDENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,

RA KINASHI H., HOPWOOD D.A.;

RT "A set of ordered cosmid and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL MOL. MICROBIOL. 21:77-96(1996).

DR EMBL: AL031184; E1314374; -.

KW TRANSMEMBRANE.

SEQUENCE 472 AA; 50203 MW; 66C4596E CRC32;

Query Match 72.0%; Score 59; DB 2; Length 472;
Best Local Similarity 70.0%; Pred. No. 3.39e+00;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 272 YLMTFSLSY 281
Oy 3 YLSTFSLDY 12

RESULT 4
ID Q20120 PRELIMINARY; PRT; 1367 AA.

AC Q20120;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-JAN-1999 (TREMBLREL. 09, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE F44G4.8 PROTEIN (EC 3.1.3.48).

GN F44G4.8.

OS CAENORHABDITIS ELEGANS.

OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;

RC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.

RN [1]
RP SEQUENCE FROM N.A.

RA THOMAS K.;

RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: 254218; E1346536; -.

DR EMBL: 249910; E1346536; JOINED.

DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.

KW HYDROLASE.

SEQUENCE 1367 AA; 152872 MW; F696EAD5 CRC32;

Query Match 72.0%; Score 59; DB 5; Length 1367;
Best Local Similarity 63.6%; Pred. No. 3.39e+00;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 160 SYLQPSFSLEY 170
Oy 2 SYLSTFSLDY 12

RESULT 5
ID Q55611 PRELIMINARY; PRT; 315 AA.

AC Q55611;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE PROTEIN-EXPORT MEMBRANE PROTEIN SECF.

GN SECF.

OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).

OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.

RN [1]
RP SEQUENCE FROM N.A.

RA TABATA S.;
 RL SUBMITTED (AUG-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCC6803;
 RX MEDLINE: 96127529.
 RA KANEKO T., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,
 RA SUGIURA M., TABATA S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. I. sequence features in the 1mb
 RT region from map positions 648 to 928 of the genome.";
 RL DNA RES. 2:153-166(1995).
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCC6803;
 RX MEDLINE: 97061201.
 RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASANIZU E., NAKAMURA Y.,
 RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAKOTO S., KIMURA T.,
 RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
 RA SHIMO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
 RA TABATA S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. PCC6803. II. Sequence determination of the entire
 RT genome and assignment of potential protein-coding regions.";
 RL DNA RES. 3:109-136(1996).
 DR EMBL: D64000; D1010770;
 SQ SEQUENCE 315 AA; 34667 MW; 4C20ED73 CRC32;
 Query Match 68.3%; Score 56; DB 2; Length 315;
 Best Local Similarity 60.0%; Pred. No. 1.12e+01;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Db 155 YLIRFOLDY 164
 11: 1 111
 3 YLSTFSLDY 12

RESULT 6
 ID 077768 PRELIMINARY; PRT; 306 AA.
 AC 077768;
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN C.
 GN HNRNPC.
 OS ORYCTOLAGUS CUNICULUS (RABBIT).
 OC EURKOTILA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 CC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
 [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=NEW ENGLAND WHITE;
 RX MEDLINE: 98438739.
 RA JIANG W., GUO X., BHAVANANDAN V.P.;
 RT "Four distinct regions in the auxiliary domain of heterogeneous
 RT nuclear ribonucleoprotein C-related proteins.";
 RL BIOCHIM. BIOPHYS. ACTA 1399:229-233(1998).
 DR EMBL: AF061582; G3660678;
 KM NUCLEOPROTEIN; RIBONUCLEOPROTEIN.
 SQ SEQUENCE 306 AA; 33684 MW; FBB28492 CRC32;
 Query Match 67.1%; Score 55; DB 6; Length 306;
 Best Local Similarity 77.8%; Pred. No. 1.65e+01;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Db 118 LSSFFLDY 126
 11: 1 111
 4 LSTFSLDY 12

RESULT 7
 ID 005237 PRELIMINARY; PRT; 357 AA.
 AC 005237;
 DT 01-JUL-1997 (TREMBLREL. 04, CREATED)

DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 39.5 KD PROTEIN.
 OS YUGH.
 OS BACILLUS SUBTILIS.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
 OC BACILLUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA ODEGA B., KONINGSSTEYN G.;
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE OF 1-214 FROM N.A.
 RC STRAIN=168;
 RA DANCHIN A.;
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA ODEGA B., KONINGSSTEYN G.;
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [4]
 RP SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE: 98044033.
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
 RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
 RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRINELL S.C., BRON S.,
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMERSON P.T.,
 RA ENTIAN K.D., ERKINGTON J., FABBET C., FERRARI E., FOULGER D.,
 RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,
 RA GHM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,
 RA GIUSEPPI G., GUY B.J., HAGA K., HALECH J., HARWOOD C.R., HENAUT A.,
 RA HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.E., ITAYA M., JONES L.,
 RA JORIS B., KARAYATA D., KASAHARA Y., KLAER-BLANCARD M., KLEIN C.,
 RA KOBAYASHI Y., KOETTER P., KONINGSSTEIN G., KROGH S., KUMANO M.,
 RA KRITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,
 RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,
 RA NOONE D., O'REILLY M., OGAWA K., OGIMAWA A., OUDGA B., PARK S.H.,
 RA PARO V., POHL T.M., PORTELELLA D., POMOLLIK S., PRESCOTT A.M.,
 RA PRESECAN E., PUTIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,
 RA RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SABAIE Y.,
 RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,
 RA SERIGUCHI J., SEKOWSKA A., SEROR S.J., SEROR P., SHIN B.S., SOLDO B.,
 RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,
 RA TAKEUCHI M., TAKAKOSHI A., TANAKA T., TERSTRA P., TOGNONI A.,
 RA TOSATO V., UCHITAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,
 RA VIARI A., WAMBUTT R., WEDLER E., WEDLER H., WEITENSGER T.,
 RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,
 RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;
 RT "The complete genome sequence of the gram-positive bacterium Bacillus
 RT subtilis".
 RL NATURE 390:249-256(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 DR EMBL: Z693934; E311527;
 DR EMBL: Z693934; E311527;
 DR PROSITE, PS00105; AA-TRANSFER_CLASS_1; 1.
 DR PFM: PF00155; aminotran_1; 1.
 KM HYPOTHETICAL PROTEIN; PYRIDOXAL PHOSPHATE.
 SQ SEQUENCE 357 AA; 39461 MW; FD6B02F9 CRC32;
 Query Match 67.1%; Score 55; DB 2; Length 357;

AC 026039;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CONSERVED HYPOTHETICAL INTEGRAL MEMBRANE PROTEIN.
GN HP1509.
OS HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).
OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; HELICOBACTER GROUP;
CC HELICOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-26695;
RX MEDLINE; 97394467.
RA TOUB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,
RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
RA LOTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLDOKER A.,
RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,
RA CORTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., MATTHEY L., WALLIN E.,
RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
RA VENTER J.C.;
RT "The complete genome sequence of the gastric pathogen *Helicobacter*
pylori [published erratum appears in *Nature* 1997 Sep
25;389(6649):412]."
RL NATURE 388:539-547(1997).
DR EMBL: AE000649; G2314690; -.
DR TIGR: HP1509; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 262 AA; 28835 MW; 5CE61FE3 CRC32;

Query Match 65.9%; Score 54; DB 2; Length 262;
Best Local Similarity 58.3%; Pred. No. 2.43e+01;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 125 AVFLSKLGLDY 136
|:|:|:|:|:|
QY 1 ASYLSTSFSLDY 12

RESULT 13
ID 065998 PRELIMINARY; PRT; 316 AA.
AC 065998;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE ORF36.
OS CLOSTRIDIUM ACETOBUTYLICUM.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;
CC CLOSTRIDIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DSM 792;
RA BEHRENS S.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U58131; G3025464; -.
SQ SEQUENCE 316 AA; 36225 MW; 720478E4 CRC32;

Query Match 65.9%; Score 54; DB 2; Length 316;
Best Local Similarity 41.7%; Pred. No. 2.43e+01;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 207 PSYLPSSGKFPDY 218
:|:|:|:|:|:|
QY 1 ASYLSTSFSLDY 12

RESULT 14
ID 020272 PRELIMINARY; PRT; 389 AA.
AC 020272;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE COSMID FA1C6.
GN FA1C6.7.
OS CAENORHABDITIS ELEGANS.
OC EURAROTIA; METAZOA; NEMATODA; SECCERNENTEA; RHABDITIA; RHABDITIDA;
CC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PEARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMER E., STRADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C.*
elegans."
RL NATURE 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA GEISEL C.;
RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RA WATERSTON R.;
RL SUBMITTED (OCT-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U39745; G1049476; -.
SQ SEQUENCE 389 AA; 44253 MW; PFE0AFC6 CRC32;

Query Match 64.6%; Score 53; DB 5; Length 389;
Best Local Similarity 70.0%; Pred. No. 3.56e+01;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

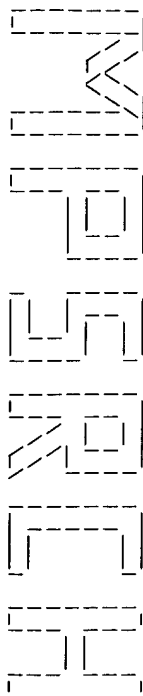
Db 157 AAYLSSSFSL 166
|:|:|:|:|:|
QY 1 ASYLSTSFSL 10

RESULT 15
ID 014411 PRELIMINARY; PRT; 783 AA.
AC 014411;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE OPT1P.
GN OPT1P.
OS CANDIDA ALBICANS (YEAST).
OC EURAROTIA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
CC CANDIDACEAE; CANDIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-1006;
RX MEDLINE; 97195785.
RA LUBKOWITZ M.A., HAUSER L., BRESLAV M., NAIDER F., BECKER J.M.;
RT "An oligopeptide transport gene from *Candida albicans*."
RL MICROBIOLOGY 143:0-0(0).
DR EMBL: U60973; G2367386; -.
SQ SEQUENCE 783 AA; 88406 MW; DC99C92E CRC32;

Query Match 64.6%; Score 53; DB 3; Length 783;
Best Local Similarity 50.0%; Pred. No. 3.56e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 406 FLSTTFSAISY 415
:|:|:|:|:|
QY 3 YLSTSFSLDY 12

Search completed: Thu Sep 2 12:40:23 1999
Job time : 30 secs.



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm

on: Thu Sep 2 12:39:25 1999; MasPar time 2.20 Seconds

Module output not generated. 154,427 Million cell updates/sec

Title: >US-08-599-226-31
Description: (1-12) from US08599226.pep
Perfect Score: 82
Sequence: 1 ASYLSTFSLDY 12

Scoring table: PAM 150
Gap 15

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot17
I:swissprot

Statistics: Mean 26.265; Variance 33.740; scale 0.778

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	63	76.8	272	1	YIGL_HAEN HYPOTHETICAL PROTEIN H	1.04e+01
2	62	75.6	1294	1	VA3B_SCHPO HYPOTHETICAL 149.2 KD	1.67e+01
3	57	69.5	296	1	CTCG_RHOSH DIHEME CYTOCHROME C-TY	1.69e+00
4	57	69.5	300	1	GP40_HUMAN PUTATIVE G PROTEIN-COU	1.69e+00
5	56	68.3	575	1	CNGX_RAT CYCLIC-NUCLEOTIDE-GATE	2.64e+00
6	55	67.1	303	1	ROC_HUMAN HETEROGENEOUS NUCLEAR	4.10e+00
7	55	67.1	438	1	SHIA_ECOLI SHIKIMATE TRANSPORT	4.10e+00
8	55	67.1	445	1	ACSC_MOOTH CORRINOID/IRON-SULFUR	4.10e+00
9	54	65.9	277	1	ICE3_CRIO APOLIPAIN PRECURSOR (EC	6.33e+00
10	54	65.9	277	1	ICE3_RAT APOLIPAIN PRECURSOR (EC	6.33e+00
11	54	65.9	277	1	ICE3_MOUSE APOLIPAIN PRECURSOR (EC	6.33e+00
12	54	65.9	611	1	YCR3_YEAST HYPOTHETICAL 69.2 KD P	6.33e+00
13	53	64.6	175	1	RL6A_YEAST 60S RIBOSOMAL PROTEIN	9.73e+00
14	53	64.6	175	1	RL6A_YEAST 60S RIBOSOMAL PROTEIN	9.73e+00
15	53	64.6	324	1	GSHB_ANASP GLUTATHIONE SYNTHETASE	9.73e+00
16	53	64.6	738	1	ST11_YEAST SERINE/THREONINE-PROTE	1.49e+01
17	52	63.4	122	1	YGT2_YEAST HYPOTHETICAL 14.4 KD P	1.49e+01
18	52	63.4	180	1	APT_MASHI ADENINE PHOSPHORIBOSYL	1.49e+01
19	52	63.4	180	1	APT_STOIO ADENINE PHOSPHORIBOSYL	1.49e+01
20	52	63.4	227	1	PESA_MYCPN CDP-DIACETYLGLYCEROL-GL	1.49e+01
21	52	63.4	306	1	CAPC_STRPN UTP-GLUCOSE-1-PHOSHA	1.49e+01
22	52	63.4	428	1	YE63_SCHPO HYPOTHETICAL 48.7 KD P	1.49e+01
23	52	63.4	475	1	YMG1_YEAST HYPOTHETICAL 55.3 KD P	1.49e+01

24	51	62.2	53	1	VG87_BPMU5 GENE 87 PROTEIN (GP87)	2.25e+01
25	51	62.2	150	1	YB9L_YEAST HYPOTHETICAL 16.0 KD P	2.25e+01
26	51	62.2	265	1	CB23_POIMU CYCLOHEXIMIDE RESISTAN	2.25e+01
27	51	62.2	552	1	YH47_CANMA CYCLOHEXIMIDE RESISTAN	2.25e+01
28	51	62.2	579	1	YH47_CABEL HYPOTHETICAL 66.0 KD P	2.25e+01
29	51	62.2	715	1	LCNC_LACIA LACTOCOCIN A TRANSPO	2.25e+01
30	51	62.2	799	1	AFSK_STRCO SERINE/THREONINE PROTE	2.25e+01
31	51	62.2	807	1	AFSK_STRCO SERINE/THREONINE PROTE	2.25e+01
32	51	62.2	817	1	TGLK_HUMAN PROTEIN-GLUTAMINE GAMM	2.25e+01
33	51	62.2	824	1	TGLK_RAT PROTEIN-GLUTAMINE GAMM	2.25e+01
34	51	62.2	836	1	TGLK_RABIT PROTEIN-GLUTAMINE GAMM	2.25e+01
35	51	62.2	3396	1	POLG_DENIS GENOME POLYPROTEIN (CO	2.25e+01
36	50	61.0	179	1	APT_HELPY ADENINE PHOSPHORIBOSYL	3.39e+01
37	50	61.0	241	1	ATP6_RHOP ATP SYNTHASE A CHAIN (3.39e+01
38	50	61.0	267	1	CB23_PETSP CHLOROPHYLL A-B BINDIN	3.39e+01
39	50	61.0	406	1	GUN1_RUMAL ENDOGLUCANASE I PRECUR	3.39e+01
40	50	61.0	409	1	GUNB_RUMAL ENDOGLUCANASE B PRECUR	3.39e+01
41	50	61.0	577	1	SYR_HAEN ARGINYL-TRNA SYNTHETAS	3.39e+01
42	50	61.0	700	1	CAOQ_RAT PRISTANOYL-COA OXIDASE	3.39e+01
43	50	61.0	835	1	FASD_ECOLI OUTER MEMBRANE USHER P	3.39e+01
44	50	61.0	844	1	MCEL_VARY MRNA CAPPING ENZYME L	3.39e+01
45	50	61.0	844	1	MCEL_VACCC MRNA CAPPING ENZYME L	3.39e+01

ALIGNMENTS

RESULT	1	STANDARD	PRT	272 AA.
ID	YIGL_HAEN			
AC	P44771			
DT	01-NOV-1995 (REL. 32, CREATED)			
DT	01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	HYPOTHETICAL PROTEIN HI0597.			
GN	HI0597.			
OS	HAEMOPHILUS INFLUENZAE.			
OC	BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURILLACEAE;			
OC	HAEMOPHILUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-RD / KW20;			
RX	MEDLINE; 95350630.			
RA	FEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,			
RA	KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,			
RA	MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOGAYNE J.D.,			
RA	SCOTT J.D., SHIRLEY R., LIU L.-I., GLODER A.A., KELLEY J.M.,			
RA	WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDDLON E., COTTON M.D.,			
RA	UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,			
RA	FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,			
RA	GREHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,			
RA	VENTER J.C.;			
RT	"whole-genome random sequencing and assembly of Haemophilus			
RT	influenzae Rd."			
RL	SCIENCE 269:496-512(1995).			
CC	-1- SIMILARITY: BELONGS TO THE COF/YBHA/YIDA/YIGL (E.COLI) / YCSE/YXEH			
CC	(B.SUBTILIS) FAMILY. STRONG. TO E.COLI YIGL.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U32741; G1573586; -			
DR	TIGR; HI0597; -			
DR	PROSITE; PS01228; COF_1; 1.			
DR	PROSITE; PS01229; COF_2; 1.			
DR	PFAM; PF00592; DUF3; 1.			
DR	HYPOTHETICAL PROTEIN.			
SO	SEQUENCE 272 AA; 30523 MW; 7F53B65C CRC32;			
Query Match	76.8%;	Score 63;	DB 1;	Length 272;


```

FT DOMAIN 1 10 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 11 31 1 (POTENTIAL).
FT DOMAIN 32 39 CYTOPLASMIC.
FT TRANSMEM 40 59 2 (POTENTIAL).
FT DOMAIN 60 71 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 72 92 3 (POTENTIAL).
FT DOMAIN 93 123 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 124 144 4 (POTENTIAL).
FT DOMAIN 145 178 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 179 197 5 (POTENTIAL).
FT DOMAIN 198 221 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 222 242 6 (POTENTIAL).
FT DOMAIN 243 258 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 259 279 7 (POTENTIAL).
FT DOMAIN 280 300 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 155 155 POTENTIAL.
SO SEQUENCE 300 AA; 31457 MW; 24383E01 CRC32;

Query Match 69.5%; Score 57; DB 1; Length 300;
Best Local Similarity 50.0%; Pred. No. 1.69e+00;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 105 YLGAAPFPLY 114
|:::|:|
3 YLSTFSFLDY 12

RESULT 5 STANDARD; PRT; 575 AA.
ID CNGX_RAT
AC 064359;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE CYCLIC-NUCLEOTIDE-GATED OLFACTORY CHANNEL OCN2 SUBUNIT.
OS RATIUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIROGNATHI; MURIDAE; MURINAE; RATIUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE; 94377458.
RA BRADLEY J., LI J., DAVIDSON N., LESTER H.A., ZINN K.;
RT "Heteromeric olfactory cyclic nucleotide-gated channels: a subunit
that confers increased sensitivity to cAMP."
RL PROC. NATL. ACAD. SCI. U.S.A. 91:8890-8894(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-OLFACTORY NEUROEPITHELIUM;
RX MEDLINE; 95000663.
RA LIMAN E.R., BUCK L.B.;
RT "A second subunit of the olfactory cyclic nucleotide-gated channel
confers high sensitivity to cAMP."
RL NEURON 13:611-621(1994).
RN [3]
RP SEQUENCE OF 7-35 FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RA BRADLEY J., ZHANG Y., BAKIN R., LESTER H.A., RONNETT G., ZINN K.;
RT SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RL -1- FUNCTION: ODORANT SIGNAL TRANSDUCTION IS PROBABLY MEDIATED
BY A G-PROTEIN COUPLED CASCADE USING CAMP AS SECOND MESSENGER.
CC THE OLFACTORY CHANNEL CAN BE SHOWN TO BE ACTIVATED BY CYCLIC
NUCLEOTIDES WHICH LEADS TO A DEPOLARIZATION OF OLFACTORY
SPENSORY NEURONS.
CC -1- SUBUNIT: HETEROOLIGOMER OF OCN1 AND OCN2 SUBUNITS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: OLFACTORY NEURONS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; U12623; G538129; -
DR EMBL; U12425; G548084; -
DR EMBL; U76219; G1753128; -
DR PROSITE; PS00888; CNMP_BINDING_1; 1.
DR PROSITE; PS00889; CNMP_BINDING_2; 1.
DR PROSITE; PS00890; CNMP_BINDING_3; 1.
DR PFAM; PF00027; CNMP_binding; 1.
DR PFAM; PF00914; CNG_membrane; 1.
KW IONIC CHANNEL; ION TRANSPORT; CAMP-BINDING; TRANSMEMBRANE;
KW MULTIGENE FAMILY; OLFACTION.
FT DOMAIN 1 33 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 34 54 1 (POTENTIAL).
FT DOMAIN 55 65 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 66 86 2 (POTENTIAL).
FT DOMAIN 87 112 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 113 133 3 (POTENTIAL).
FT DOMAIN 134 168 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 169 189 4 (POTENTIAL).
FT DOMAIN 190 244 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 245 265 5 (POTENTIAL).
FT DOMAIN 266 400 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 401 421 6 (POTENTIAL).
FT DOMAIN 422 575 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 356 500 CAMP (BY SIMILARITY).
FT BINDING 415 415 CAMP (POTENTIAL).
FT BINDING 430 430 CAMP (POTENTIAL).
SO SEQUENCE 575 AA; 65674 MW; C545A708 CRC32;

Query Match 68.3%; Score 56; DB 1; Length 575;
Best Local Similarity 63.6%; Pred. No. 2.64e+00;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 64 SYLVAMPFLDY 74
|:::|:|
2 YLSTFSFLDY 12

RESULT 6 STANDARD; PRT; 303 AA.
ID ROC_HUMAN
AC P07910; P22628;
DT 01-AUG-1988 (REL. 08, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEINS C1/C2 (HNRNP C1 AND HNRNP
C2).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE; 90093350.
RA BURD C.G., SWANSON M.S., GOERLACH M., DREYFUS G.;
RT "Primary structures of the heterogeneous nuclear ribonucleoprotein
A2, B1, and C2 proteins: a diversity of RNA binding proteins is
generated by small peptide inserts."
RL PROC. NATL. ACAD. SCI. U.S.A. 86:9788-9792(1989).
RN [2]
RP SEQUENCE OF 1-107 AND 121-303 FROM N.A.
RC MEDLINE; 87257872.
RA SWANSON M.S., NAKAGAWA T.Y., LEVAN K., DREYFUS G.;
RT "Primary structure of human nuclear ribonucleoprotein particle C
proteins: conservation of sequence and domain structures in
heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins."
RL MOL. CELL. BIOL. 7:1731-1739(1987).
RN [3]
RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE; 90067819.
RA MERRILL B.M., BARNETT S.F., LESTOURGEON W.M., WILLIAMS K.R.;
RT "Primary structure differences between proteins C1 and C2 of HeLa 40S
nuclear ribonucleoprotein particles."
RL NUCLEIC ACIDS RES. 17:8441-8449(1989).

```

[4]
RN STRUCTURE BY NMR OF 1-94.
RX MEDLINE: 92329450.
RA WITTEKING M., GOERLACH M., FRIEDRICH M., DREYFUS G., MUELLER L.;
RT "1H, 13C, and 15N NMR assignments and global folding pattern of the
RL RNA-binding domain of the human hnRNP C proteins.";
RN BIOCHEMISTRY 31:6254-6265(1992).
[5]
RN STRUCTURE BY NMR OF 1-94.
RX MEDLINE: 92371436.
RA GOERLACH M., WITTEKING M., BECKMAN R.A., MUELLER L., DREYFUS G.;
RT "Interaction of the RNA-binding domain of the hnRNP C proteins with
RL RNA.";
RN EMBO J. 11:3289-3295(1992).
CC -1- FUNCTION: MAY PLAY A ROLE IN RIBONUCLEOSOME ASSEMBLY BY
CC NEUTRALIZING BASIC PROTEINS SUCH AS A AND B CORE HNRP.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; COMPONENT OF RIBONUCLEOSOMES.
CC -1- PTM: PHOSPHORYLATED (PROBABLE).
CC -1- ALTERNATIVE PRODUCTS: C1 AND C2 PROTEINS PROBABLY RESULT FROM
CC ALTERNATIVE SPLICING OF THE PRE-MRNA. C2 IS IDENTICAL TO C1
CC EXCEPT FOR AN INSERT OF 13 AA (AA 108-120) ONLY PRESENT IN C2.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RNP).
CC -1- SIMILARITY: HIGH, TO X LAEYIS PROTEIN C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M29063; G337455; -
DR EMBL: M16342; G306875; -
DR PIR: C34504; C34504.
DR PIR: A26885; A26885.
DR AARHUS/GENE-2DPAGE: 7207; IEF.
DR AARHUS/GENE-2DPAGE: 7222; IEF.
DR MIM: 164020; -
DR SWISS-2DPAGE: P07910; HUMAN.
DR PROSITE: PS00030; RNP_1; 1.
DR PFAM: PF00076; rtm: 1
KM NUCLEAR PROTEIN; RNA-BINDING; RIBONUCLEOPROTEIN; PHOSPHORYLATION;
KM ALTERNATIVE SPLICING.
FT DOMAIN 50 57 RNA-BINDING (RNP1).
FT DOMAIN 155 161 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 181 303 ASP/GLU-RICH (ACIDIC).
FT VARSPLIC 108 120 MISSING (IN HNRP C1).
FT SEQUENCE 303 AA; 33298 MW; 09615140 CRC32;
Query Match 67.1%; Score 55; DB 1; Length 303;
Best Local Similarity 77.8%; Pred. No. 4.10e+00;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 118 LSSSFLDY 126
QY 4 LSTFSFLDY 12
RESULT 7
ID SHIA_ECOLI STANDARD; PRT; 438 AA.
AC P76350;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE SHIKIMATE TRANSPORTER.
GN SHIA.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;

RX MEDLINE: 98192527.
RA WHIPP M.J., CAMAKARIS H., PITTARD A.J.;
RT "Cloning and analysis of the shia gene, which encodes the shikimate
RL transport system of Escherichia coli K-12.";
RN GENE 209:185-192(1998).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE: 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GODDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL SCIENCE 277:1453-1474(1997).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN-K12.
RX MEDLINE: 97251358.
RA ITOH T., AIBA H., BABU T., FUJITA K., HAYASHI K., INADA T.,
RA ISONO K., KASAI H., KIMURA S., KITAGAWA M., KITAGAWA M.,
RA MAKINO K., MIKI T., MIZOBUCHI K., MORI H., MORI T., MOTOKURA K.,
RA NAKADE S., NAKAMURA Y., NASHIMOTO H., NISHIO Y., OSHIMA T.,
RA SATO N., SANEI G., SEKI Y., SIVASUNDARAM S., TAGAMI H.,
RA TAKEDA J., TAKEMOTO K., WADA C., YAMAMOTO Y., HORIUCHI T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RL corresponding to the 40.1-50.0 mln region on the linkage map.";
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U88529; G1850982; -
DR EMBL: AE000290; G1788292; -
DR EMBL: D90837; G1736645; -
DR ECOGENE: B620205; SHIA.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
DR PFAM: PF00083; sugar_tr: 1.
KM TRANSPORT; TRANSMEMBRANE; INNER MEMBRANE; SYMPORT.
FT TRANSMEM 28 48 POTENTIAL.
FT TRANSMEM 64 84 POTENTIAL.
FT TRANSMEM 109 129 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 193 213 POTENTIAL.
FT TRANSMEM 225 275 POTENTIAL.
FT TRANSMEM 287 307 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 387 407 POTENTIAL.
FT TRANSMEM 411 431 POTENTIAL.
SQ SEQUENCE 438 AA; 47817 MW; 8773D0DD CRC32;
Query Match 67.1%; Score 55; DB 1; Length 438;
Best Local Similarity 50.0%; Pred. No. 4.10e+00;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 267 YIYTAFALNV 276
QY 3 YLSTFSFLDY 12
RESULT 8
ID ACSC_MOOTH STANDARD; PRT; 445 AA.
AC 007340;
DT 01-FEB-1995 (REL. 31, CREATED)

DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CORRIDOID/IRON-SULFUR PROTEIN, LARGE SUBUNIT (C/FE-SP).
GN AGSC.
OS MOORELLA THERMOACETICA (CLOSTRIDIUM THERMOACETICUM).
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOORELLA GROUP;
CC MOORELLA.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE: 93194857.
RA LU W.-P., SCHIAU I., CUNNINGHAM J.R., RAGSDALE S.W.;
RT "Sequence and expression of the gene encoding the
corridoid/iron-sulfur protein from Clostridium thermoaceticum and
RT reconstitution of the recombinant protein to full activity";
RL J. BIOL. CHEM. 268:5605-5614(1993).
RN [2]
RP SEQUENCE OF 1-7 FROM N.A., AND SEQUENCE OF 1-18.
RC STRAIN-DSM 521;
RX MEDLINE: 89098907.
RA ROBERTS D.L., JAMES-HAGSTROM J.E., GARVIN D.K., GORST C.M.,
RONDONIUST J.A., BAUR J.R., HAASE F.C., RAGSDALE S.W.;
RT "Cloning and expression of the gene cluster encoding key proteins
involved in acetyl-CoA synthesis in Clostridium thermoaceticum: CO
dehydrogenase, the corridoid/Fe-S protein, and methyltransferase";
RL PROC. NATL. ACAD. SCI. U.S.A. 86:32-36(1989).
CC -1- FUNCTION: ACTS AS A METHYL GROUP CARRIER IN THE ANAEROBIC ACETYL-
COA PATHWAY (WOOD/LJUNGDAHL PATHWAY) OF CARBON MONOXIDE AND CARBON
DIOXIDE FIXATION.
CC COA PATHWAY: HETERODIMER OF A LARGE AND SMALL CHAIN.
CC -1- SUBUNIT: HETERODIMER OF A LARGE AND SMALL CHAIN.
CC -----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L07099; G144854; -.
DR PIR: A46621; A46621.
DR HSP: P01088; 1BFA.
KW COBALT; IRON-SULFUR; 4FE-4S; CARBON DIOXIDE FIXATION.
FT INIT_MET 0
FT METAL 16 IRON-SULFUR (4FE-4S) (PROBABLE).
FT METAL 19 IRON-SULFUR (4FE-4S) (PROBABLE).
FT METAL 24 IRON-SULFUR (4FE-4S) (PROBABLE).
SQ SEQUENCE 445 AA; 48021 MW; F6ACB6C6 CRC32;
Query Match 67.1%; Score 55; DB 1; Length 445;
Best Local Similarity 60.0%; Pred. No. 4,10e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 337 YVTNFSFLTY 346
QY 3 YLSTFSFLDY 12

RX MEDLINE: 96183185.
RA WANG X., ZELENISKI N.G., YANG J., SAKAI J., BROWN M.S.,
RA GOLDSTEIN J.L.;
RT "Cleavage of sterol regulatory element binding proteins (SREBPs) by
RT CPP32 during apoptosis";
RL EMBO J. 15:1012-1020(1996).
CC -1- FUNCTION: INVOLVED IN THE ACTIVATION CASCADE OF CASPASES
CC RESPONSIBLE FOR APOPTOSIS EXECUTION. AT THE ONSET OF APOPTOSIS IT
CC PROTEOLYTICALLY CLEAVES POLY(ADP-RIBOSE) POLYMERASE (PARP) AT A
CC 216-ASP-1-GLY-217 BOND. CLEAVES AND ACTIVATES SREBP. REGULATORY
CC ELEMENT BINDING PROTEINS (SREBPs) BETWEEN THE BASIC HELIX-LOOP-
CC CLEAVES AND ACTIVATES CASPASE-6, -7 AND -9 (BY SIMILARITY).
CC -1- SUBUNIT: HETERODIMER OF A 17 KD (P17) AND A 12 KD (P12) SUBUNIT
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- PTM: CLEAVAGE BY GRANZYME B, CASPASE-6, -8 AND -10 GENERATES THE
CC TWO ACTIVE SUBUNITS. ADDITIONAL PROCESSING OF THE PROPEPTIDES IS
CC LIKELY DUE TO THE AUTOCATALYTIC ACTIVITY OF THE ACTIVATED
CC PROTEASE. ACTIVE HETERODIMERS BETWEEN THE SMALL SUBUNIT OF
CC CASPASE-7 PROTEASE AND THE LARGE SUBUNIT OF CPP32 ALSO OCCUR AND
CC VICE VERSA (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C14; ALSO KNOWN AS THE
CC CASPASE FAMILY.
CC -----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U27463; G124444; -.
DR PROSITE: PS01121; CASPASE_HIS; 1.
DR PROSITE: PS01122; CASPASE_CYS; 1.
DR PFAM: PF00655; ICE_P10; 1.
DR PFAM: PF00656; ICE_P20; 1.
DR HSP: P42574; 1PAD.
KW HYDROLASE; THIOL PROTEASE; ZMOGEN; APOPTOSIS.
FT PROPEP 1
FT PROPEP 9
FT CHAIN 29 175 APOPAIN P17 SUBUNIT.
FT CHAIN 176 277 APOPAIN P12 SUBUNIT.
FT ACT_SITE 121 121 BY SIMILARITY.
FT ACT_SITE 163 163 BY SIMILARITY.
SQ SEQUENCE 277 AA; 31612 MW; 29FID5A7 CRC32;
Query Match 65.9%; Score 54; DB 1; Length 277;
Best Local Similarity 50.0%; Pred. No. 6,33e+00;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Db 32 YLDSYKMDY 41
QY 3 YLSTFSFLDY 12

RESULT 9
ID ICE3_CRILIO STANDARD: PRT: 277 AA.
AC 060431;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE APOPAIN PRECURSOR (EC 3.4.22.-) (CYSTEINE PROTEASE CPP32) (YAMA
DE PROTEIN) (CPP-32) (CASPASE-3) (SREBP CLEAVAGE ACTIVITY 1)
DE (SCA-1).
GN CASP3 OR CPP32.
OS CARCINUS LONGICAUDATUS (LONG-TAILED HAMSTER) (CHINESE HAMSTER).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
OC RODENTIA: SCUROGNATHI: MORIDAE: CRICETINAE: CRICETULUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;

RESULT 10
ID ICE3_RAT STANDARD: PRT: 277 AA.
AC P55213; P70543; Q62993; P97699;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE APOPAIN PRECURSOR (EC 3.4.22.-) (CYSTEINE PROTEASE CPP32) (YAMA
DE PROTEIN) (CPP-32) (CASPASE-3) (SREBP CLEAVAGE ACTIVITY 1)
DE (SCA-1) (LICE) (IRP).
GN CASP3 OR CPP32.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
OC RODENTIA: SCUROGNATHI: MORIDAE: MORINAE: RATTUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96358624.

RA JUAN T.S.-C., MCNIECE I.K., JENKINS N.A., GILBERT D.J., COPELAND N.G.,
 RA FLETCHER F.A.:
 RT "Molecular characterization of mouse and rat CPP32 beta gene encoding
 RT a cysteine protease resembling interleukin-1 beta converting enzyme
 RT and CED-3.";
 RL ONCOGENE 13:749-755(1996).
 RM [2]
 RP SEQUENCE OF 30-241 FROM N.A.
 RC TISSUE-OVARY;
 RX MEDLINE: 96042508.
 RA FLAWS J.A., KUGU K., TROVICH A.M., DESANTI A., TILLY K.I.,
 RA HIRSHFIELD A.N., TILLY J.L.:
 RT "Interleukin-1 beta-converting enzyme-related proteases (IRPs) and
 RT mammalian cell death: dissociation of IRP-induced oligonucleosomal
 RT endonuclease activity from morphological apoptosis in granulosa cells
 RT of the ovarian follicle.";
 RL ENDOCRINOLOGY 136:5042-5053(1995).
 RM [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE: 97184204.
 RA NI B., WU X., DU Y., SU Y., HAMILTON-BYRD E., ROCKEY P.K.,
 RA ROSECK P. JR., POIRIER G.G., PAUL S.M.:
 RT "Cloning and expression of a rat brain interleukin-1beta-converting
 RT enzyme (ICE)-related protease (IRP) and its possible role in
 RT apoptosis of cultured cerebellar granule neurons.";
 RL J. NEUROSCI. 17:1561-1569(1997).
 RM [4]
 RP SEQUENCE OF 1-264 FROM N.A.
 RC YAKOVLEV A.G.:
 RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: INVOLVED IN THE ACTIVATION CASCADE OF CASPASES
 CC RESPONSIBLE FOR APOPTOSIS EXECUTION. AT THE ONSET OF APOPTOSIS IT
 CC PROTEOLYTICALLY CLEAVES POLY(ADP-RIBOSE) POLYMERASE (PARP) AT A
 CC 216-ASP-1-GLY-217 BOND. CLEAVES AND ACTIVATES STEROL REGULATOR
 CC ELEMENT BINDING PROTEINS (SREBPS) BETWEEN THE BASIC HELIX-LOOP-
 CC HELIX LEUCINE ZIPPER DOMAIN AND THE MEMBRANE ATTACHMENT DOMAIN.
 CC CLEAVES AND ACTIVATES CASPASE-6, -7 AND -9 (BY SIMILARITY).
 CC -1- SUBUNIT: HETERODIMER OF A 17 KD (P17) AND A 12 KD (P12) SUBUNIT
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, LIVER, AND MUSCLE
 CC BUT NOT IN KIDNEY OR TESTIS.
 CC -1- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED IN NEURON-ENRICHED REGIONS
 CC OF THE DEVELOPING BRAIN, BUT DOWN-REGULATED TO LOW LEVELS IN THE
 CC ADULT BRAIN.
 CC -1- PTM: CLEAVAGE BY GRANTZYME B, CASPASE-6, -8 AND -10 GENERATES THE
 CC TWO ACTIVE SUBUNTS. ADDITIONAL PROCESSING OF THE PROPEPTIDES IS
 CC LIKELY DUE TO THE AUTOCATALYTIC ACTIVITY OF THE ACTIVATED
 CC PROTEASE. ACTIVE HETERODIMERS BETWEEN THE SMALL SUBUNIT OF
 CC CASPASE-7 PROTEASE AND THE LARGE SUBUNIT OF CPP32 ALSO OCCUR AND
 CC VICE VERSA (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C14; ALSO KNOWN AS THE
 CC CASPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL: U49930; G1518936; -
 DR EMBL: U34685; G1004371; -
 DR EMBL: U84410; G1814027; -
 DR EMBL: U58656; G181643; -
 DR PROSITE: PS01121; CASPASE_HIS; 1.
 DR PFAM: PF00655; ICE_P10; 1.
 DR PFAM: PF00656; ICE_P20; 1.
 DR HSSP: P43574; 1PAU.
 KM HYDROLASE; THIOLE PROTEASE; ZYMOGEN; APOPTOSIS.

FT PROPEP 1 9 BY SIMILARITY.
 FT PROPEP 10 28 BY SIMILARITY.
 FT CHAIN 29 175 APOPAIN P17 SUBUNIT.
 FT CHAIN 176 277 APOPAIN P12 SUBUNIT.
 FT ACT_SITE 121 121 BY SIMILARITY.
 FT ACT_SITE 163 163 BY SIMILARITY.
 FT ACT_SITE 163 163 BY SIMILARITY.
 FT CONFLICT 25 29 KSMDS -> OVD (IN REF. 4).
 FT CONFLICT 170 170 C -> S (IN REF. 2).
 FT CONFLICT 178 178 T -> A (IN REF. 2).
 FT CONFLICT 182 182 M -> V (IN REF. 2).
 FT CONFLICT 187 187 I -> K (IN REF. 2).
 FT CONFLICT 190 187 E -> G (IN REF. 3).
 FT CONFLICT 199 199 T -> S (IN REF. 2).
 FT CONFLICT 211 211 D -> G (IN REF. 2).
 FT CONFLICT 236 236 L -> I (IN REF. 4).
 FT CONFLICT 245 245 T -> M (IN REF. 3).
 SQ SEQUENCE 277 AA; 31491 MM; 0E106140 CRC32;
 Query Match 65.9%; Score 54; DB 1; Length 277;
 Best Local Similarity 50.0%; Pred. No. 6,338+00;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 DB 32 YLDSYKMDY 41
 Oy 3 YLSTSFSDY 12
 RESULT 11
 AC ICE3_MOUSE STANDARD: PRT: 277 AA.
 ID P70677; 008668;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE APOPAIN PRECURSOR (BC 3.4.22.-) (CYSTEINE PROTEASE CPP32) (YAMA
 DE PROTEIN) (CSP-32) (CASPASE-3) (SREBP CLEAVAGE ACTIVITY 1)
 DE (SCA-1) (LICE).
 GN CASP3 OR CPP32.
 OS MUS MUSCULUS (MUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIROGNATHI; MURIDAE; MURINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96358624.
 RA JUAN T.S.-C., MCNIECE I.K., JENKINS N.A., GILBERT D.J., COPELAND N.G.,
 RA FLETCHER F.A.:
 RT "Molecular characterization of mouse and rat CPP32 beta gene encoding
 RT a cysteine protease resembling interleukin-1 beta converting enzyme
 RT and CED-3.";
 RL ONCOGENE 13:749-755(1996).
 RM [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97224429.
 RA MURASA T., URASE K., MOMOI M.Y., KIMURA I., MOMOI T.:
 RT "Specific expression of CPP32 in sensory neurons of mouse embryos and
 RT activation of CPP32 in the apoptosis induced by a withdrawal of
 RT NGF.";
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 231:770-774(1997).
 RM [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C3H/AN;
 RC MEDLINE: 97190206.
 RA VAN DE CRAEN M., VANDENABELE P., DECLERCQ W., VAN DEN BRANDE I.,
 RA VAN DE LOO G., MOLEMAN F., SCHOTTE P., VAN CIEKINGE W., BEYERT R.,
 RA FIERS W.:
 RT "Characterization of seven murine caspase family members.";
 RL FEBS LETT. 403:61-69(1997).
 RM [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA FERNANDES-ALMEIDA T., LITWACK G., ALNEMRI E.S.:
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: INVOLVED IN THE ACTIVATION CASCADE OF CASPASES
 CC RESPONSIBLE FOR APOPTOSIS EXECUTION. AT THE ONSET OF APOPTOSIS IT

CC PROTEOLYTICALLY CLEAVES POLY(ADP-RIBOSE) POLYMERASE (PARP) AT A
 CC 216-ASP-1-GLY-217 BOND. CLEAVES AND ACTIVATES STEROL REGULATORY
 CC ELEMENT BINDING PROTEINS (SRBPS) BETWEEN THE BASIC HELIX-LOOP-
 CC HELIX LEUCINE ZIPPER DOMAIN AND THE MEMBRANE ATTACHMENT DOMAIN.
 CC CLEAVES AND ACTIVATES CASPASE-6, -7 AND -9 (BY SIMILARITY).
 CC CLEAVES IL-1 BETA BETWEEN AN ASP AND AN ALA, RELEASING THE MATURE
 CC CYTOKINE WHICH IS INVOLVED IN A VARIETY OF INFLAMMATORY PROCESSES.
 CC -1 SUBUNIT: HETERODIMER OF A 17 KD (P17) AND A 12 KD (P12) SUBUNIT
 CC (BY SIMILARITY).
 CC -1 SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1 PPM: CLEAVAGE BY GRANZYME B, CASPASE-6, -8 AND -10 GENERATES THE
 CC TWO ACTIVE SUBUNITS. ADDITIONAL PROCESSING OF THE PROPEPTIDES IS
 CC LIKELY DUE TO THE AUTOCATALYTIC ACTIVITY OF THE ACTIVATED
 CC PROTEASE. ACTIVE HETERODIMERS BETWEEN THE SMALL SUBUNIT OF
 CC CASPASE-7 PROTEASE AND THE LARGE SUBUNIT OF CPP32 ALSO OCCUR AND
 CC VICE VERSA (BY SIMILARITY).
 CC -1 TISSUE SPECIFICITY: HIGHEST EXPRESSION IN SPLEEN, LUNG, LIVER,
 CC KIDNEY AND HEART. LOWER EXPRESSION IN BRAIN, SKELETAL MUSCLE AND
 CC TESTIS.
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY C14; ALSO KNOWN AS THE
 CC CASPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U54803; G1518968; -.
 CC EMBL; U54802; G1518968; JOINED.
 CC EMBL; U49829; G1518934; -.
 CC EMBL; D86352; G1945544; -.
 CC EMBL; D86352; D1022582; -.
 CC EMBL; Y13086; E315505; -.
 CC EMBL; U19522; G2114328; -.
 CC MGD; MGI:107739; CASP3.
 CC PROSITE; PS01121; CASPASE_HIS; 1.
 CC PROSITE; PS01122; CASPASE_CYS; 1.
 CC DR PFAM; PF00655; ICE_P10; 1.
 CC DR PFAM; PF00656; ICE_P20; 1.
 CC DR HSSP; P42574; 1PAU.
 CC KM HYDROLASE: THIOL PROTEASE: ZYMOGEN: APOPTOSIS.
 CC FT PROPEP 1 9
 CC FT PROPEP 10 28
 CC FT CHAIN 29 175
 CC CHAIN 176 277
 CC ACT_SITE 121 121
 CC ACT_SITE 163 163
 CC ACT_SITE 17 17
 CC CONFLICT 51 51
 CC CONFLICT 84 84
 CC CONFLICT 95 95
 CC CONFLICT 97 97
 CC CONFLICT 128 128
 CC CONFLICT 135 135
 CC SEQUENCE 277 AA; 31474 MW; EA7DCLF CRC32;
 CC -----
 CC Query Match 65.9%; Score 54; DB 1; Length 277;
 CC Best Local Similarity 50.0%; Pred. No. 6.33e+00;
 CC Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 69.2 KD PROTEIN IN HSP30-PMP1 INTERGENIC REGION.
 GN YCR023C OR YCR23C OR YCR241.
 OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 CC EUKARYOTA. FUNGI: ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
 CC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC RX MEDLINE; 92245758.
 CC RA BOLLÉ P. A., GILLIQUET V., BERBEN G., DUMONT J., HILGER F.;
 CC RT "The complete sequence of K3B, a 7.9 kb fragment between PCK1 and
 CC RT CRY1 on chromosome III, reveals the presence of seven open reading
 CC RT frames.";
 CC RL YEAST 8:205-213(1992).
 CC CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
 CC CC -1 SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X59720; E264487; -.
 CC DR PIR; S19434; S19434.
 CC DR PIR; S22273; S22273.
 CC DR PIR; S22273; S22273.
 CC KM HYPOTHETICAL PROTEIN: TRANSPORT; TRANSMEMBRANE.
 CC FT TRANSMEM 90 110
 CC FT TRANSMEM 153 173
 CC FT TRANSMEM 200 220
 CC FT TRANSMEM 354 372
 CC FT TRANSMEM 414 434
 CC FT TRANSMEM 443 463
 CC FT TRANSMEM 543 563
 CC SEQUENCE 611 AA; 69198 MW; 74CDBDE CRC32;
 CC -----
 CC Query Match 65.9%; Score 54; DB 1; Length 611;
 CC Best Local Similarity 60.0%; Pred. No. 6.33e+00;
 CC Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

RA MARIS E., MENEZES S., MILLER N., NHAN M., PAULEY A., PELUSO D.,
 RA RIFKEN L., RILES L., TAICH A., TREVASKIS E., VIGNATI D.,
 RA WILCOX L., WOLDMAN P., VAUDIN M., WILSON R., WATERSTON R.,
 RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE OF 1-24.
 RA OTAKA E., HIGO K.-I., ITOH T.,
 RT "Yeast ribosomal proteins: VIII. Isolation of two proteins and
 RT sequence characterization of twenty-four proteins from cytoplasmic
 RT ribosomes."
 RL MOL. GEN. GENET. 195:544-546(1984).
 CC -1- THERE ARE TWO GENES FOR L6 IN YEAST.
 CC -1- SIMILARITY: BELONGS TO THE L6E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D10226; G218511; -
 DR EMBL: U22382; G217063; -
 DR PIR: S11257; S11257.
 DR PIR: S28945; S28945.
 DR SGD: L0001714; RPL6B.
 DR PROSITE: PS01170; RIBOSOMAL_L6E; 1.
 DR PFAM: PF01159; L6e; 1.
 KW RIBOSOMAL PROTEIN; MULTIGENE FAMILY.
 FT INIT_MET 0
 FT CONFLICT 0
 SQ SEQUENCE 175 AA; 19855 MW; EADAB462 CRC32;
 P -> L (IN REF. 1).
 Query Match 64.6%; Score 53; DB 1; Length 175;
 Best Local Similarity 87.5%; Pred. No. 9.73e+00;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Db 157 YLSTFSL 164
 Y 3 YLSTFSL 10
 RESULT 14
 ID RLG6_YEAST STANDARD; PRT; 175 AA.
 AC Q02326;
 DT 01-OCT-1993 (REL. 27, CREATED)
 DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 RL 60S RIBOSOMAL PROTEIN L6-A (L17) (YL16) (RPL8).
 RL RPL6A OR YL16A OR YML073C.
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 OS EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 93003323.
 RA HASHIMOTO T., SUZUKI K., MIZUTA K., OTAKA E.,
 RT "Yeast ribosomal proteins: XIV. Complete nucleotide sequences of the
 RT two genes encoding Saccharomyces cerevisiae YL16."
 RL BIOCHIM. BIOPHYS. ACTA 1132:195-198(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA BROWN D., BOWMAN S., BARRELL B.G., RAJANDREAM M.A.;
 RL SUBMITTED (OCT-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- THERE ARE TWO GENES FOR L6 IN YEAST.
 CC -1- SIMILARITY: BELONGS TO THE L6E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D10225; G218509; -
 DR EMBL: Z46373; G914877; -
 DR PIR: S28944; S28944.
 DR SGD: L0004165; RPL6A.
 DR PROSITE: PS01170; RIBOSOMAL_L6E; 1.
 DR PFAM: PF01159; L6e; 1.
 KW RIBOSOMAL PROTEIN; MULTIGENE FAMILY.
 FT INIT_MET 0
 FT CONFLICT 0
 SQ SEQUENCE 175 AA; 19830 MW; 2E155789 CRC32;
 Query Match 64.6%; Score 53; DB 1; Length 175;
 Best Local Similarity 87.5%; Pred. No. 9.73e+00;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Db 157 YLSTFSL 164
 Y 3 YLSTFSL 10
 RESULT 15
 ID GSHB_ANASP STANDARD; PRT; 324 AA.
 AC P45480; Q43879;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 RL GLUTATHIONE SYNTHETASE (EC 6.3.2.3) (GLUTATHIONE SYNTHASE) (GSH
 DE SYNTHETASE) (GSH-S).
 GN GSHB OR GSH-II.
 OS ANABAENA SP. (STRAIN PCC 7120).
 OC BACTERIA; CYANOBACTERIA; NOSTOCALES; NOSTOCACEAE; ANABAENA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96001250.
 RA DOHERTY H.M., ADAMS D.G.;
 RT "Cloning and sequence of ftsz and flanking regions from the
 RT cyanobacterium Anabaena PCC 7120."
 RL GENE 163:93-96(1995).
 RN [2]
 RP SEQUENCE OF 24-324 FROM N.A.
 RX MEDLINE: 96099685.
 RA ZHANG C.C., HUGUENIN S., FRIY A.;
 RT "Analysis of genes encoding the cell division protein ftsz and a
 RT glutathione synthetase homologue in the cyanobacterium Anabaena sp.
 RL PCC 7120".
 RL RES. MICROBIOL. 146:445-455(1995).
 CC -1- CATALYTIC ACTIVITY: ATP + GAMMA-L-GLUTAMYL-L-CYSTEINE + GLYCINE -
 CC ADP + ORTHOPHOSPHATE + GLUTATHIONE.
 CC -1- PATHWAY: SECOND STEP IN GLUTATHIONE BIOSYNTHESIS.
 CC -1- SIMILARITY: TO OTHER PROKARYOTIC GSH SYNTHASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U14408; G555916; -
 DR EMBL: Z31371; E190151; -
 DR HSSP: P04425; ZGLT.
 KW GLUTATHIONE BIOSYNTHESIS; LIGASE; ATP-BINDING.
 FT CONFLICT 50
 FT CONFLICT 50
 FT CONFLICT 243
 SQ SEQUENCE 324 AA; 35870 MW; 61CCDD40 CRC32;
 Query Match 64.6%; Score 53; DB 1; Length 324;
 Best Local Similarity 45.8%; Pred. No. 9.73e+00;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 102 AYLXATYVLDY 112
: | : : : | | |
OY 2 SYLSTSFSLDY 12

Search completed: Thu Sep 2 12:39:34 1999
Job time : 9 secs.

THIS PAGE BLANK (USPTO)

RESULT	ENTRY	TITLE
2	S58307	#type complete hypothetical protein SPAC18B1.11 - fission yeast (Schizosaccharomyces pombe)

ORGANISM #formal_name Schizosaccharomyces pombe
DATE 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
ACCESSIONS S58307
REFERENCE S58307
#authors Devlin, K.; Churcher, C.M.
#submission Submitted to the EMBL Data Library, August 1995
#accession S58307
#molecule_type DNA
#residues 1-1294 #label DEV
#cross-references EMBL:Z50728; NID:G929886; PID:G929897
GENETICS
#map_position 1L
SUMMARY #length 1294 #molecular-weight 149191 #checksum 7509
Query Match 75.6%; Score 62; DB 2; Length 1294;
Best Local Similarity 72.7%; Pred. No. 5,45e-01;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 456 SYDTSFSLDY 466
||:|||||
QY 2 YLSTFSLDY 12

RESULT 3
ENTRY JC5714 #type complete
TITLE G protein-coupled receptor 40 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 03-Dec-1997 #sequence_revision 23-Jan-1998 #text_change
17-Mar-1999
ACCESSIONS JC5714
REFERENCE JC5714
#authors Sawzdargo, M.; George, S.R.; Nguyen, T.; Xu, S.; Kolakowski
Jr., L.F.; O'Dowd, B.F.
#journal Biochem. Biophys. Res. Commun. (1997) 239:543-547
#title A cluster of four novel human G protein-coupled receptor
genes occurring in close proximity to CD22 gene on
chromosome 19q13.1.
#cross-references MUID:98008875
#accession JC5714
#status nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-300 #label SAN
#cross-references GB:AF024687; NID:G2612945; PID:G2612946
KEYWORDS glycoprotein; lipoprotein; thiolester bond
FEATURE
13-32 #domain transmembrane #status predicted #label TM1
43-64 #domain transmembrane #status predicted #label TM2
83-102 #domain transmembrane #status predicted #label TM3
125-144 #domain transmembrane #status predicted #label TM4
184-206 #domain transmembrane #status predicted #label TM5
222-243 #domain transmembrane #status predicted #label TM6
260-280 #domain transmembrane #status predicted #label TM7
155,165 #binding_site carbohydrate (Asn) (covalent) #status
predicted
289 #binding_site palmitate (Cys) (covalent) #status
predicted
SUMMARY #length 300 #molecular-weight 31457 #checksum 7805
Query Match 69.5%; Score 57; DB 2; Length 300;
Best Local Similarity 50.0%; Pred. No. 4,54e+00;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 105 YLGAAPFLGY 114
||::|||
QY 3 YLSTFSLDY 12

RESULT 4
ENTRY S76267 #type complete
TITLE hypothetical protein - Synecocystis sp. (strain PCC 6803)
ORGANISM #formal_name Synecocystis sp.
#variety PCC 6803

DATE 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change
21-Aug-1998
ACCESSIONS S76267
REFERENCE S74322
#authors Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.;
Nakamura, Y.; Miyajima, N.; Hirosewa, M.; Sugita, M.;
Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.;
Murai, A.; Nakazaki, N.; Natsu, K.; Okumura, S.; Shimpou,
S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;
Yasuda, M.; Tabata, S.
#journal DNA Res. (1996) 3:109-136
#title Sequence analysis of the genome of the unicellular
cyanobacterium Synecocystis sp. PCC6803. II. Sequence
determination of the entire genome and assignment of
potential protein-coding regions.
#cross-references MUID:97061201
#accession S76267
#status preliminary
#molecule_type DNA
#residues 1-315 #label KAN
#cross-references EMBL:D64000; GB:AB001339; NID:G1001484; PID:G1010770;
PID:G1001494
#note the nucleotide sequence was submitted to the EMBL Data
Library, June 1996
SUMMARY #length 315 #molecular-weight 34667 #checksum 4918
Query Match 68.3%; Score 56; DB 2; Length 315;
Best Local Similarity 60.0%; Pred. No. 6,85e+00;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Db 155 YLTFPOLDY 164
||:||||
QY 3 YLSTFSLDY 12

RESULT 5
ENTRY I59327 #type complete
TITLE olfactory cyclic nucleotide-gated cation channel - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
25-Apr-1997
ACCESSIONS I59327; I58165
REFERENCE I59327
#authors Bradley, J.; Li, J.; Davidson, N.; Lester, H.A.; Zinn, K.
#journal Proc. Natl. Acad. Sci. U.S.A. (1994) 91:8890-8894
#title Heteromeric olfactory cyclic nucleotide-gated channels: A new
subunit that confers increased sensitivity to cAMP.
#cross-references MUID:94377458
#accession I59327
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-575 #label RES
#cross-references EMBL:U12623; NID:G558128; PID:G558129
REFERENCE I58165
#authors Liman, E.R.; Buck, L.B.
#journal Neuron (1994) 13:611-621
#title A second subunit of the olfactory cyclic nucleotide-gated
channel confers high sensitivity to cAMP.
#cross-references MUID:95000663
#accession I58165
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-575 #label RE2
#cross-references EMBL:U12425; NID:G548083; PID:G548084
CLASSIFICATION #superfamily cAMP receptor protein cyclic nucleotide-binding
domain homology
FEATURE
348-472
SUMMARY #domain cAMP receptor protein cyclic nucleotide-binding
domain homology #label CA2
#length 575 #molecular-weight 65674 #checksum 1183
Query Match 68.3%; Score 56; DB 2; Length 575;
Best Local Similarity 63.6%; Pred. No. 6,85e+00;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 64 SYLVAMEVLDY 74
||| : |||
QY 2 SYLSTFSLDY 12

RESULT 6
ENTRY C34504 #type complete
TITLE heterogeneous ribonucleic acid particle protein C2 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 10-Sep-1997

ACCESSIONS
REFERENCE C34504
#authors Burd, C.G.; Swanson, M.S.; Goerlach, M.; Dreyfuss, G.
#journal Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9788-9792
#title Primary structures of the heterogeneous nuclear ribonucleoprotein A2, B1, and C2 proteins: a diversity of RNA binding proteins is generated by small peptide inserts.

#cross-references MUID:9009350
#accession C34504
#status preliminary
#molecule_type mRNA
#residues 1-303 #label BUR
#cross-references GB:M29063; NID:G337454; PID:G337455
CLASSIFICATION #superfamily_unassigned.ribonucleoprotein_repeat-containing proteins; ribonucleoprotein_repeat_homology alternative_splicing

KEYWORDS
FEATURE 17-77
SUMMARY #domain.ribonucleoprotein_repeat_homology #label RRM1
#length 303 #molecular_weight 33298 #checksum 1186

Query Match 67.1%; Score 55; DB 2; Length 303;
Best Local Similarity 77.8%; Pred. NO. 1.03e+01;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 118 LSSSFDLY 126
||| : |||
QY 4 LSTFSLDY 12

RESULT 7
ENTRY G70010 #type complete
TITLE aspartate aminotransferase homolog yugh - Bacillus subtilis
ORGANISM #formal_name Bacillus subtilis
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1998

ACCESSIONS
REFERENCE G70010
#authors Kunet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans, A.; Braun, M.; Briganti, S.C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denicot, F.; Devine, K.M.; Duesterhoeft, A.; Ehrlich, S.D.; Emerson, P.T.; Enlian, K.D.; Errington, J.; Fribet, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.; Giuseppe, G.; Guy, B.J.; Haga, K.; Halach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.; Katsuhara, Y.; Klier-Blanchard, M.; Klein, C.; Kobayashi, Y.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Porwolik, S.; Prescott, A.M.; Presecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.;

Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scanlon, E.; Schleich, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, S.J.; Seror, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenbol, M.; Vannier, F.; Vassart, A.; Viari, A.; Wambutt, R.; Wedler, E.; Wedler, H.; Weitzengger, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

#journal Nature (1997) 390:249-256
#title The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

#cross-references MUID:98044033
#accession G70010
#status preliminary; nucleic acid sequence not shown; translation not shown

#molecule_type DNA
#residues 1-357 #label KUN
#cross-references GB:Z99120; GB:AL009126; NID:G2635613; PID:ell184218; PID:G2635636
##experimental_source strain 168

GENETICS
#gene yugh
CLASSIFICATION #superfamily aspartate transaminase
SUMMARY #length 357 #molecular_weight 39461 #checksum 3672

Query Match 67.1%; Score 55; DB 2; Length 357;
Best Local Similarity 50.0%; Pred. NO. 1.03e+01;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Db 75 SRYLSNFDLY 86
||| : |||
QY 1 ASYLSTFSLDY 12

RESULT 8
ENTRY G64962 #type complete
TITLE shikimate transport protein shiA - Escherichia coli
ORGANISM #formal_name Escherichia coli
DATE 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Feb-1999

ACCESSIONS
REFERENCE G64962; S78630
#authors Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.

#journal Science (1997) 277:1453-1462
#title The complete genome sequence of Escherichia coli K-12.
#cross-references MUID:97426617
#accession G64962

#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-438 #label BLAT
#cross-references GB:AE000290; GB:U00096; NID:G2367125; PID:G1788292; UNCP:01981

##experimental_source strain K-12, substrain MG1655

REFERENCE S78630
#authors Whipp, M.J.; Camakaris, H.; Pittard, A.J.
#journal Gene (1998) 209:185-192
#title Cloning and analysis of the shiA gene, which encodes the shikimate transport system of Escherichia coli K-12.

#accession S78630
#status nucleic acid sequence not shown
#molecule_type DNA
#residues 1-438 #label WHI
#cross-references EMBL:U88529

GENETICS
#gene shiA

```

#FUNCTION
#description involved in the uptake of shikimate
#pathway aromatic amino acid biosynthesis
#CLASSIFICATION #superfamily citrate utilization determinant
#KEYWORDS transmembrane protein; transport protein
FEATURE
32-50 #domain transmembrane #status predicted #label TM1\
58-76 #domain transmembrane #status predicted #label TM2\
96-114 #domain transmembrane #status predicted #label TM3\
117-135 #domain transmembrane #status predicted #label TM4\
171-189 #domain transmembrane #status predicted #label TM5\
198-216 #domain transmembrane #status predicted #label TM6\
260-278 #domain transmembrane #status predicted #label TM7\
291-309 #domain transmembrane #status predicted #label TM8\
319-337 #domain transmembrane #status predicted #label TM9\
340-358 #domain transmembrane #status predicted #label TM10\
390-408 #domain transmembrane #status predicted #label TM11\
414-432 #domain transmembrane #status predicted #label TM12\
SUMMARY #length 438 #molecular-weight 47817 #checksum 9704

Query Match 67.1%; Score 55; DB 2; Length 438;
Best Local Similarity 50.0%; Pred. NO. 1.03e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 267 YVTAFALNY 276
1:1:1:1:1
Oy 3 YLSTFSFLDY 12

RESULT 9
ENTRY #type complete
TITLE cortinoid/iron-sulfur protein large chain - Clostridium
#ALTERNATE_NAMES C/Fe-Sp large chain AcSC
#ORGANISM #formal_name Clostridium thermaceticum
#DATE 21-Sep-1993 #sequence_revision 04-Oct-1996 #text_change
18-Sep-1998
ACCESSIONS A46621
REFERENCE A46621
#authors Lu, W.P.; Schiau, I.; Cunningham, J.R.; Ragsdale, S.W.
#journal J. Biol. Chem. (1993) 268:5605-5614
#title Sequence and expression of the gene encoding the
cortinoid/iron-sulfur protein from Clostridium
thermaceticum and reconstitution of the recombinant
protein to full activity.
#cross-references MUID:93194857
#accession A46621
##molecule_type DNA; protein
##residues 1-446 ##label LUI
#note Sequence extracted from NCBI backbone (NCBIN:127877,
NCBIP:127878)

GENETICS acsc
#gene #superfamily cortinoid/iron-sulfur protein large chain
#CLASSIFICATION 4Fe-4S; carbon dioxide fixation; electron transfer;
#KEYWORDS heterodimer; iron-sulfur protein; metalloprotein
FEATURE
2-446 #product cortinoid/iron-sulfur protein large chain
#status experimental #label MT\
#binding_site 4Fe-4S cluster (Cys) (covalent) #status
predicted
SUMMARY #length 446 #molecular-weight 48153 #checksum 3949

Query Match 67.1%; Score 55; DB 1; Length 446;
Best Local Similarity 60.0%; Pred. NO. 1.03e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 338 YVTNFSFLTY 347
1:1:1:1:1
Oy 3 YLSTFSFLDY 12

RESULT 10

```

```

ENTRY      557545      #type complete
TITLE      probable membrane protein YPR012w - yeast (Saccharomyces
ALTERNATE_NAMES      hypothetical protein YP951.05
ORGANISM    cerevisiae
DATE        10-Oct-1995 #format_name Saccharomyces cerevisiae
                21-Nov-1997 #format_name Saccharomyces cerevisiae
                03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change
                21-Nov-1997 #sequence_revision 03-Nov-1995 #text_change

ACCESSIONS      557545
REFERENCE      557541
                Bowman, S.
                #authors
                #submission
                #accession      557545
                #molecule_type DNA
                ##residues      1-84 ##label BOW
                ##cross-references EMBL:Z49919; NID:g887584; PID:g887589; MIPS:YPR012w
                ##experimental_source strain AB972

GENETICS
#map_position      16R
KEYWORDS      transmembrane protein
FEATURES
FEATURE_1      11-27      #domain transmembrane #status predicted #label TM
SUMMARY      #length 84 #molecular_weight 9904 #checksum 7108

Query Match      65.9%: Score 54; DB 2; Length 84;
Best Local Similarity 60.0%: Pred. No. 1,53e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db      6 YLNTAFSLAY 15
        ||: |||| |
Oy      3 YLSTFSLDY 12

RESULT      11
ENTRY      567699      #type complete
TITLE      probable membrane protein YDL151c - yeast (Saccharomyces
ALTERNATE_NAMES      hypothetical protein D1554
ORGANISM    cerevisiae
DATE        12-Jul-1996 #format_name Saccharomyces cerevisiae
                14-Nov-1997 #format_name Saccharomyces cerevisiae
                12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change

ACCESSIONS      567699
REFERENCE      567693
                Pera, J.; Blugeon, C.; Delavau, T.; Jacq, C.
                #authors
                #submission
                #accession      567699
                #molecule_type DNA
                ##residues      1-193 ##label PER
                ##cross-references EMBL:Z74199; NID:g1431236; PID:e253062; PID:g1431237;
                MIPS:YDL151c
                ##experimental_source strain S286C

GENETICS
#map_position      4L
KEYWORDS      transmembrane protein
FEATURES
FEATURE_1      114-130      #domain transmembrane #status predicted #label TM
SUMMARY      #length 193 #molecular_weight 20893 #checksum 4697

Query Match      65.9%: Score 54; DB 2; Length 193;
Best Local Similarity 40.0%: Pred. No. 1,53e+01;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Db      149 FLATAFGLNF 158
        :|:|:|:|:|:
Oy      3 YLSTFSLDY 12

RESULT      12
ENTRY      167437      #type fragment
TITLE      cysteine protease (EC 3.4.22.) p32 - rat (fragment)
ORGANISM    #format_name Rattus norvegicus #common_name Norway rat
DATE        26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
                17-Mar-1999 #sequence_revision 26-Jul-1996 #text_change

ACCESSIONS      167437

```


REFERENCE
#authors 153300
#journal K.I.: Hirschfeld, A.N.: Tilly, J.L.
#title Endocrinology (1995) 136:5042-5053
Interleukin-1 beta-converting enzyme-related proteases (IRPs)
and mammalian cell death: dissociation of IRP-induced
oligonucleosomal endonuclease activity from morphological
apoptosis in granulosa cells of the ovarian follicle.
#cross-references MUID:96042508
#accession 167437
#status preliminary: translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-212 ##label RES
#cross-references EMBL:U34685; NID:g1004370; PID:g1004371
KEYWORDS
#keywords Cysteine proteinase: hydrolase
#length 212 #checksum 472
SUMMARY
Query Match 65.9%; Score 54; DB 2; Length 212;
Best Local Similarity 50.0%; Pred. No. 1.53e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
OY 3 YLSTSFSLDY 12
11 : 11 : 11 :
11 : 11 : 11 :
OY 3 YLSTSFSLDY 12

RESULT 13
ENTRY B71811 #type complete
#title hypotetical protein jhp1402 - Helicobacter pylori (strain
J99)
#formal_name Helicobacter pylori
#strain J99
#date 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
05-Mar-1999

ORGANISM
#variety strain J99

DATE 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
05-Mar-1999

ACCESSIONS
#accession B71811
#authors Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.;
Dolg, P.C.; Smith, D.R.; Noonan, B.; Guild, B.C.; deJonghe,
B.L.; Carmel, G.; Tummino, P.J.; Caruso, A.;
Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.;
Merberg, D.; Mills, S.D.; Jiang, O.; Taylor, D.E.; Vovis,
G.F.; Trust, T.J.
#journal Nature (1999) 397:176-180
#title Genomic sequence comparison of two unrelated isolates of the
human gastric pathogen Helicobacter pylori.
#cross-references MUID:99120557
#accession B71811
#status Preliminary
#molecule_type DNA
#residues 1-220 ##label ARN
#cross-references GB:AE001562; GB:AE001439; NID:g4156017; PID:g4156023
#experimental_source strain J99

GENETICS
#gene jhp1402
#superfamily Escherichia coli YgiH protein
SUMMARY
#length 220 #molecular-weight 23804 #checksum 9738

Query Match 65.9%; Score 54; DB 2; Length 220;
Best Local Similarity 58.3%; Pred. No. 1.53e+01;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 ASLSTSFSLDY 12

DB 83 AVLSKLFGLDY 94
11 : 11 : 11 :
11 : 11 : 11 :
OY 1 ASLSTSFSLDY 12

RESULT 14
ENTRY E64708 #type complete
#title conserved hypothetical integral membrane protein HP1509 -
Helicobacter pylori (strain 26695)
#formal_name Helicobacter pylori
#date 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change
13-Nov-1998

ACCESSIONS
REFERENCE E64708
#authors A64520
#journal Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.;
Sutton, G.G.; Fleischmann, R.D.; Ketchum, K.A.; Klenk,
H.P.; Gill, S.; Dougherty, B.A.; Nelson, K.; Quackenbush,
J.; Zhou, L.; Kirness, E.F.; Peterson, S.; Loftus, B.;
Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.;
McKenney, K.; Fitzgerald, L.M.; Lee, N.; Adams, M.D.;
Hickey, E.K.; Berg, D.E.; Gocayne, J.D.; Uterback, T.R.;
Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.;
Fujii, C.; Bowman, C.; Wathey, L.; Wallin, E.; Hayes,
W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
C.M.; Venter, J.C.
#journal Nature (1997) 388:539-547
#title The complete genome sequence of the gastric pathogen
Helicobacter pylori.
#cross-references MUID:97394467
#accession E64708
#status preliminary: nucleic acid sequence not shown;
translational not shown

CLASSIFICATION
#superfamily Escherichia coli YgiH protein
SUMMARY
#length 262 #molecular-weight 28835 #checksum 768

Query Match 65.9%; Score 54; DB 2; Length 262;
Best Local Similarity 58.3%; Pred. No. 1.53e+01;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 ASLSTSFSLDY 12

DB 125 AVLSKLFGLDY 136
11 : 11 : 11 :
11 : 11 : 11 :
OY 1 ASLSTSFSLDY 12

RESULT 15
ENTRY JC5410 #type complete
#title Cpp32 protein - mouse
#formal_name Mus musculus #common_name house mouse
#date 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change
17-Mar-1999

ACCESSIONS
#accession JC5410
#authors Mukasa, T.; Urase, K.; Momoi, M.Y.; Kimura, I.; Momoi, T.
#journal Biochem. Biophys. Res. Commun. (1997) 231:770-774
#title Specific expression of Cpp32 in sensory neurons of mouse
embryos and activation of Cpp32 in the apoptosis induced by
a withdrawal of NGF.
#cross-references MUID:97224425
#accession JC5410
#status nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-277 ##label MUX
#cross-references DBJ:D86352
#experimental_source embryo

COMMENT
#comment This protein is involved in the apoptosis of dorsal root ganglia
neurons.
SUMMARY
#length 277 #molecular-weight 31392 #checksum 7414

Query Match 65.9%; Score 54; DB 2; Length 277;
Best Local Similarity 50.0%; Pred. No. 1.53e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
OY 3 YLSTSFSLDY 12

DB 32 YLSTSFSLDY 41
11 : 11 : 11 :
11 : 11 : 11 :
OY 3 YLSTSFSLDY 12

Search completed: Thu Sep 2 12:39:07 1999
Job time : 15 secs.

THIS PAGE BLANK (USPTO)

M P S R E L F
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch.p protein - protein database search, using Smith-Waterman algorithm

on: Thu Sep 2 12:40:39 1999: Maspar time 1.39 Seconds
87.779 Million cell updates/sec
Molecular output not generated.

Title: >US-08-599-226-31
Description: (1-12) from US08599226.pep
Perfect Score: 82
Sequence: 1 ASYLSTSFSLDY 12

Scoring table: PAM 150
Gap 15

Searched: 106580 seqs, 10152877 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PC19_COMB 4:backfiles1

Statistics: Mean 16.747; Variance 61.471; scale 0.272

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	54	65.9	277	2	US-08-890-Sequence 2, Applicatio	6.26e+01
2	52	63.4	306	3	PCT-US95-0-Sequence 13, Applicati	9.53e+01
3	49	59.8	267	2	US-08-570-Sequence 2, Applicati	1.78e+02
4	49	59.8	898	2	US-08-693-Sequence 36, Applicati	1.78e+02
5	49	59.8	908	2	US-08-588-Sequence 3, Applicatio	1.78e+02
6	49	59.8	908	2	US-08-693-Sequence 33, Applicati	1.78e+02
7	49	59.8	960	2	US-08-693-Sequence 8, Applicatio	1.78e+02
8	49	59.8	960	2	US-08-588-Sequence 3, Applicatio	1.78e+02
9	49	59.8	960	2	US-08-355-Sequence 8, Applicatio	1.78e+02
10	48	58.5	117	3	PCT-US96-0-Sequence 5, Applicatio	2.18e+02
11	48	58.5	117	3	PCT-US96-0-Sequence 11, Applicati	2.18e+02
12	48	58.5	329	1	US-08-348-Sequence 12, Applicati	2.18e+02
13	48	58.5	343	1	US-08-348-Sequence 10, Applicati	2.18e+02
14	48	58.5	969	1	US-07-671-Sequence 4, Applicati	2.18e+02
15	48	58.5	1174	3	PCT-US95-0-Sequence 29, Applicati	2.18e+02
16	48	58.5	1174	2	US-08-639-Sequence 29, Applicati	2.18e+02
17	48	58.5	1174	1	US-08-349-Sequence 29, Applicati	2.18e+02
18	48	58.5	1174	1	US-08-349-Sequence 28, Applicati	2.18e+02
19	48	58.5	1174	2	US-08-598-Sequence 29, Applicati	2.18e+02
20	48	58.5	1177	4	US-07-828-Patent No. 5169629	2.18e+02
21	48	58.5	1177	4	US-07-828-Patent No. 5169629	2.18e+02
22	48	58.5	1177	3	PCT-US92-1-Sequence 8, Applicatio	2.18e+02
23	48	58.5	1177	1	US-07-920-Sequence 2, Applicatio	2.18e+02

24	48	58.5	1178	4	5254799-5	Patent No. 5254799.	2.18e+02
25	48	58.5	1178	1	US-08-446-Sequence 5, Applicatio	2.18e+02	2.18e+02
26	48	58.5	1178	2	US-08-463-Sequence 5, Applicati	2.18e+02	2.18e+02
27	48	58.5	1182	2	US-08-598-Sequence 34, Applicati	2.18e+02	2.18e+02
28	48	58.5	1182	1	US-08-349-Sequence 34, Applicati	2.18e+02	2.18e+02
29	48	58.5	1188	3	PCT-US95-0-Sequence 34, Applicati	2.18e+02	2.18e+02
30	48	58.5	1188	2	US-08-239-Sequence 34, Applicati	2.18e+02	2.18e+02
31	48	58.5	1188	2	US-08-639-Sequence 34, Applicati	2.18e+02	2.18e+02
32	47	57.3	180	2	US-08-447-Sequence 8, Applicatio	2.67e+02	2.67e+02
33	47	57.3	180	2	US-08-447-Sequence 7, Applicatio	2.67e+02	2.67e+02
34	47	57.3	192	2	US-08-531-Sequence 29, Applicati	2.67e+02	2.67e+02
35	47	57.3	279	3	PCT-US95-0-Sequence 5, Applicatio	2.67e+02	2.67e+02
36	47	57.3	281	2	US-08-810-Sequence 2, Applicatio	2.67e+02	2.67e+02
37	47	57.3	282	1	US-08-118-Sequence 52, Applicati	2.67e+02	2.67e+02
38	47	57.3	282	3	PCT-US93-0-Sequence 52, Applicati	2.67e+02	2.67e+02
39	47	57.3	325	4	5320941-2	Patent No. 5320941.	2.67e+02
40	47	57.3	485	1	US-08-068-Sequence 1, Applicatio	2.67e+02	2.67e+02
41	47	57.3	485	1	US-08-464-Sequence 1, Applicatio	2.67e+02	2.67e+02
42	47	57.3	505	1	US-08-068-Sequence 3, Applicatio	2.67e+02	2.67e+02
43	47	57.3	505	1	US-08-464-Sequence 3, Applicatio	2.67e+02	2.67e+02
44	47	57.3	700	3	PCT-US95-0-Sequence 2, Applicatio	2.67e+02	2.67e+02
45	47	57.3	1487	2	US-08-760-Sequence 2, Applicatio	2.67e+02	2.67e+02

ALIGNMENTS

RESULT	1	STANDARD:	PRT:	277 AA.
ID	US-08-890-542A-2			
XX	xxxxxx			
XX				
DT				
XX				
DE	Sequence 2, Application US/08890542A			
CC	Patent No. 5840509			
CC	GENERAL INFORMATION:			
CC	APPLICANT: NI, Binhui			
CC	APPLICANT: Paul, Steven M			
CC	APPLICANT: Wu, Xin			
CC	TITLE OF INVENTION: PROTEASE AND RELATED NUCLEIC ACID			
CC	TITLE OF INVENTION: COMPOUNDS			
CC	NUMBER OF SEQUENCES: 3			
CC	CORRESPONDENCE ADDRESS:			
CC	ADDRESSEE: Eli Lilly and Company			
CC	STREET: Lilly Corporate Center			
CC	CITY: Indianapolis			
CC	STATE: Indiana			
CC	COUNTRY: United States of America			
CC	ZIP: 46285			
CC	COMPUTER READABLE FORM:			
CC	MEDIUM TYPE: Floppy disk			
CC	COMPUTER: IBM PC compatible			
CC	OPERATING SYSTEM: PC-DOS/MS-DOS			
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30			
CC	CURRENT APPLICATION DATA:			
CC	APPLICATION NUMBER: US/08/890,542A			
CC	FILING DATE: 09-JUL-1997			
CC	CLASSIFICATION: 435			
CC	ATTORNEY/AGENT INFORMATION:			
CC	NAME: Gaylo, Paul J.			
CC	REGISTRATION NUMBER: 36,808			
CC	REFERENCE/DOCKET NUMBER: X-10704			
CC	TELECOMMUNICATION INFORMATION:			
CC	TELEPHONE: (317) 276-0756			
CC	TELEFAX: (317) 276-3861			
CC	INFORMATION FOR SEQ ID NO: 2:			
CC	SEQUENCE CHARACTERISTICS:			
CC	LENGTH: 277 amino acids			
CC	TYPE: amino acid			
CC	TOPOLOGY: linear			
CC	MOLECULE TYPE: protein			

SQ SEQUENCE 277 AA; 31449 MW; 406172 CN;

Query Match	65.98;	Score 54;	DB 2;	Length 277;
Best Local Similarity	50.08;	Pred. No. 6.26e+01;		
Matches	5;	Conservative	3;	Mismatches 2;
				Indels 0;
				Gaps 0;

```
Db      32 YLDSSYKMDY 41
        ||:|: ||
QY      3 YLSTSFSLDY 12
```

RESULT	2	STANDARD;	PRT;	306 AA
ID	PCT-US95-06119-13			
XX				

Sequence 13, Application PC/TUS9506119
Sequence 13, Application PC/TUS9506119
GENERAL INFORMATION:

CC TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE CAPSULAR
CC TITLE OF INVENTION: POLYSACCHARIDE GENES AND FLANKING REGIONS
CC NUMBER OF SEQUENCES: 20
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Arnold, White & Durkee
CC STREET: P.O. Box 4433
CC City: Houston
CC STATE: TX
CC COUNTRY: United States of America
CC Zip: 77210

CC
CC MEDIA TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MC-DOS/ASCII
CC SOFTWARE: PatentIn Release #1.0, Version
CC SOFTWARE: #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/06119
CC FILING DATE: CONCURRENTLY HERewith
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/243,546
CC FILING DATE: 16-MAY-1994
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:

Query Match	63.48;	Score 52;	DB 3;	Length 306;
Best Local Similarity	40.08;	Pred. No. 9.53e+01;		
Matches	4;	Conservative	2;	Indels 0; Gaps 0

```

Db      67  YEDSTFELEY  76
      | : : | | : |
QY      3  YLSTSFSLDY  12

```

RESULT	3		
ID	US-08-570-929-2	STANDARD;	PRT; 267 AA

Sequence 2, Application US/08570929
Sequence 2, Application US/08570929
Patent No. 573503

CC TITLE OF INVENTION: Anti-sense RNA for CAB Transcript to Reduce
CC
CC TITLE OF INVENTION: Chlorophyll Content in Plants
CC
CC NUMBER OF SEQUENCES: 2
CC
CC COMPUTER READABLE FORM:
CC
CC MEDIUM TYPE: Floppy disk
CC
CC COMPUTER: IBM PC compatible
CC
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC
CC SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CC
CC CURRENT APPLICATION DATA:
CC
CC APPLICATION NUMBER: US/08/570,929
CC

Query Match	59.88;	Score 49;	DB 2;	Length 267;
Best Local Similarity	50.0%;	Pred. No. 1.78e+02;		
Matches	6;	Conservative	3;	Mismatches 3; Indels 0; Gaps 0

```

Db      67 PSYLTGEFPGDY 78
      :|||: | : ||
QY      1 ASYLSTSFSLDY 12

```

RESULT	4		
ID	US-08-693-697-36	STANDARD;	PRT; 898 AA
XX			
AC	xxxxxx		

Sequence 36, Application US/08693697
Sequence 36, Application US/08693697
Patent No. 5669610

CC APPLICANT: Snodgrass, H. R.
CC APPLICANT: Ciofffi, Joseph
CC APPLICANT: Zupancic, Thomas J.
CC APPLICANT: Shafer, Alan W.
CC TITLE OF INVENTION: Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN
CC TITLE OF INVENTION: RECEPTOR
CC NUMBER OF SEQUENCES: 38
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Penille & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: FastSeq for Windows Version 2.0b

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/693,697
CC FILING DATE: 05-AUG-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Poissant, Brian M.
CC REGISTRATION NUMBER: 28,462
CC REFERENCE/DOCKET NUMBER: 8907-0037-999
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 650-493-4935
CC TELEFAX: 650-493-5556
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 36:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 898 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC FRAGMENT TYPE: internal
CC SEQUENCE 898 AA; 102731 MW; 4634755 CN;
Query Match 59.8%; Score 49; DB 2; Length 898;
Best Local Similarity 50.0%; Pred. No. 1.78e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Db 19 YVITAFNLSY 28
OY 3 YLSTFSFLDY 12
RESULT 5
ID US-08-588-526-3 STANDARD: PRT; 908 AA.
AC xxxxxx
XX
XX
XX
XX
DE Sequence 3, Application US/08588526
XX
CC Sequence 3, Application US/08588526
CC Patent No. 5882860
CC GENERAL INFORMATION:
CC APPLICANT: Snodgrass, H.
CC APPLICANT: Cioffi, Joseph
CC APPLICANT: Zupancic, Thomas
CC APPLICANT: Shafer, Alan
CC TITLE OF INVENTION: DETECTION OF A LEPTIN RECEPTOR
CC TITLE OF INVENTION: VARIANT
CC TITLE OF INVENTION: AND METHODS FOR REGULATING OBESITY
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: US
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/588,526
CC FILING DATE: 18-JAN-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Poissant, Brian M.
CC REGISTRATION NUMBER: 28,462
CC REFERENCE/DOCKET NUMBER: 8907-030
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090

CC TELEFAX: (212) 869-9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 908 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 908 AA; 103856 MW; 4758599 CN;
Query Match 59.8%; Score 49; DB 2; Length 908;
Best Local Similarity 50.0%; Pred. No. 1.78e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Db 19 YVITAFNLSY 28
OY 3 YLSTFSFLDY 12
RESULT 6
ID US-08-693-697-33 STANDARD: PRT; 908 AA.
AC xxxxxx
XX
XX
XX
XX
DE Sequence 33, Application US/08693697
XX
CC Sequence 33, Application US/08693697
CC Patent No. 5869610
CC GENERAL INFORMATION:
CC APPLICANT: Snodgrass, H. R.
CC APPLICANT: Cioffi, Joseph
CC APPLICANT: Zupancic, Thomas J.
CC APPLICANT: Shafer, Alan W.
CC TITLE OF INVENTION: Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN
CC TITLE OF INVENTION: RECEPTOR
CC NUMBER OF SEQUENCES: 38
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: FastSeq for Windows Version 2.0b
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/693,697
CC FILING DATE: 05-AUG-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Poissant, Brian M.
CC REGISTRATION NUMBER: 28,462
CC REFERENCE/DOCKET NUMBER: 8907-0037-999
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 650-493-4935
CC TELEFAX: 650-493-5556
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 33:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 908 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC FRAGMENT TYPE: internal
CC SEQUENCE 908 AA; 103702 MW; 4753903 CN;
Query Match 59.8%; Score 49; DB 2; Length 908;

Best Local Similarity 50.0%; Pred. No. 1.78e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 19 YVITAFNLISY 28
|:|:|:|
QY 3 YLSTFSLDY 12

RESULT 7
ID US-08-693-697-8 STANDARD; PRT: 960 AA.
XX xxxxxx

DE Sequence 8, Application US/08693697

CC Sequence 8, Application US/08693697

CC Patent No. 5869610

CC GENERAL INFORMATION:

CC APPLICANT: Snodgrass, H. R.

CC APPLICANT: Cioffi, Joseph

CC APPLICANT: Zupancic, Thomas J.

CC APPLICANT: Shafer, Alan W.

CC TITLE OF INVENTION: Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN

CC NUMBER OF SEQUENCES: 38

CC CORRESPONDENCE ADDRESS:

CC STREET: 1155 Avenue of the Americas

CC CITY: New York

CC STATE: New York

CC COUNTRY: USA

CC ZIP: 10036-2711

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: FASTSEQ for Windows Version 2.0b

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/693.697

CC FILING DATE: 05-AUG-1996

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Poissant, Brian M.

CC REGISTRATION NUMBER: 28,462

CC TELEPHONE: 650-493-4935

CC TELEFAX: 650-493-5556

CC MOLECULE TYPE: protein

CC TOPOLOGY: linear

CC TYPE: amino acid

CC LENGTH: 960 amino acids

CC SEQUENCE 960 AA: 109634 MW: 5296019 CN:

CC Query Match 59.8%; Score 49; DB 2; Length 960;

CC Best Local Similarity 50.0%; Pred. No. 1.78e+02;

CC Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 19 YVITAFNLISY 28
|:|:|:|
QY 3 YLSTFSLDY 12

RESULT 8
ID US-08-588-190-3 STANDARD; PRT: 960 AA.
XX xxxxxx

AC xxxxxx
CC xxxxxx

DT Sequence 3, Application US/08588190

DE Sequence 3, Application US/08588190

CC Patent No. 5856098

CC GENERAL INFORMATION:

CC APPLICANT: Snodgrass, H. Ralph

CC APPLICANT: Cioffi, Joseph

CC APPLICANT: Zupancic, Thomas Joel

CC APPLICANT: Shafer, Alan Wayne

CC TITLE OF INVENTION: DETECTION OF A LEPTIN RECEPTOR

CC NUMBER OF SEQUENCES: 4

CC CORRESPONDENCE ADDRESS:

CC STREET: 1155 Avenue of The Americas

CC CITY: New York

CC STATE: NY

CC COUNTRY: USA

CC ZIP: 10036-2811

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Diskette

CC OPERATING SYSTEM: DOS

CC SOFTWARE: FASTSEQ Version 2.0

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/588.190

CC FILING DATE: 18-JAN-1996

CC CLASSIFICATION: 435

CC PRIOR APPLICATION NUMBER:

CC FILING DATE:

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Poissant, Brian M.

CC REGISTRATION NUMBER: 28,462

CC TELEPHONE/DOCKET NUMBER: 008907-0029-999

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 650-493-4935

CC TELEFAX: 650-493-5556

CC MOLECULE TYPE: protein

CC TOPOLOGY: linear

CC TYPE: amino acid

CC STRANDEDNESS: single

CC SEQUENCE 960 AA: 109698 MW: 5295408 CN:

CC Query Match 59.8%; Score 49; DB 2; Length 960;

CC Best Local Similarity 50.0%; Pred. No. 1.78e+02;

CC Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 19 YVITAFNLISY 28
|:|:|:|
QY 3 YLSTFSLDY 12

RESULT 9
ID US-08-355-888A-8 STANDARD; PRT: 960 AA.
XX xxxxxx

DE Sequence 8, Application US/08355888A

CC Sequence 8, Application US/08355888A

CC Patent No. 5763211

CC GENERAL INFORMATION:

CC APPLICANT: Snodgrass, H. R.

CC APPLICANT: Cioffi, Joseph

```
CC APPLICANT: Zupancic, Thomas J.  
CC APPLICANT: Shafer, Alan W.  
CC TITLE OF INVENTION: Hu-BI-219, A NOVEL HUMAN HEMATOPOIETIN  
CC TITLE OF INVENTION: RECEPTOR  
CC NUMBER OF SEQUENCES: 31  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10036-2711  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/355,888A  
CC FILING DATE: 14-DEC-1994  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Poissant, Brian M.  
CC REGISTRATION NUMBER: 28,462  
CC REFERENCE/DOCKET NUMBER: 7225-078  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-9741/8864  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 8:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 960 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 960 AA; 109634 MW; 5296019 CN;  
SQ  
  
Query Match 59.8%; Score 49; DB 2; Length 960;  
Best Local Similarity 50.0%; Pred. No. 1.78e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
Db 19 YVITAFNLSTY 28  
|:|:|:|  
Oy 3 YLSTFSLSLY 12  
  
RESULT 10  
STD PCT-US96-08730-5 STANDARD; PRT; 37 AA.  
XXXXXX  
XX  
DT  
DE Sequence 5, Application PC/TUS9608730  
XX  
CC GENERAL INFORMATION:  
CC APPLICANT: Cassels, Frederick  
CC APPLICANT: Anderson, Jeffrey  
CC APPLICANT: Carter, John Mark  
CC TITLE OF INVENTION: Methods of Raising Antibodies Against E.  
CC TITLE OF INVENTION: Coil of the Family CSF-CFA./1  
CC NUMBER OF SEQUENCES: 15  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Glenna Hendicks  
CC STREET: P.O. Box 2509  
CC CITY: Fairfax  
CC STATE: VA  
CC COUNTRY: USA  
CC ZIP: 22031  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible
```

CC	OPERATING SYSTEM: PC-DOS/MS-DOS
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: PCT/US96/08730
CC	FILING DATE: 03-JUN-1996
CC	CLASSIFICATION:
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: Hendricks, Glenna
CC	REGISTRATION NUMBER: 32,535
CC	REFERENCE/DOCKET NUMBER: PCT/US96/08730
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: (703) 591-4470
CC	TELEFAX: (703) 591-4428
CC	INFORMATION FOR SEQ ID NO: 5:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 37 amino acids
CC	TYPE: amino acid
CC	STRANDEDNESS: single
CC	TOPOLOGY: unknown
CC	MOLECULE TYPE: peptide
CC	HYPOTHEICAL: NO
CC	ANTI-SENSE: NO
CC	FRAGMENT TYPE: Internal
SQ	SEQUENCE 37 AA: 3864 MW: 7776 CN:
Db	Query Match 58.5%; Score 48; DB 3; Length 37;
	Best Local Similarity 50.0%; Pred.No. 2.18e+02;
Qy	Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
	1 ASYSTSFSLDY 12
	: :
Db	22 GSYLEPLAVELTY 33
Qy	1 ASYSTSFSLDY 12
RESULT 11	
ID	PCT-US96-08730-11 STANDARD; PRT: 117 AA.
AC	xxxxxx
XX	
DT	
XX	
DE	Sequence 11, Application PC/TUS9608730
XX	
CC	Sequence 11, Application PC/TUS9608730
CC	GENERAL INFORMATION:
CC	APPLICANT: Cassels, Frederick
CC	APPLICANT: Anderson, Jeffrey
CC	APPLICANT: Carter, John Mark
CC	TITLE OF INVENTION: Methods of Raising Antibodies Against E.
CC	TITLE OF INVENTION: Coll of the Family CSF-CFA./1
CC	NUMBER OF SEQUENCES: 15
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: Glenna Hendricks
CC	STREET: P.O. Box 2509
CC	CITY: Fairfax
CC	STATE: VA
CC	COUNTRY: USA
CC	ZIP: 22031
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Floppy disk
CC	COMPUTER: IBM PC compatible
CC	OPERATING SYSTEM: PC-DOS/MS-DOS
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: PCT/US96/08730
CC	FILING DATE: 03-JUN-1996
CC	CLASSIFICATION:
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: Hendricks, Glenna
CC	REGISTRATION NUMBER: 32,535
CC	REFERENCE/DOCKET NUMBER: PCT/US96/08730
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: (703) 591-4470

CC TELEFAX: (703) 591-4428
CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 117 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: peptide
CC HYPOTHEICAL: NO
CC ANTI-SENSE: NO
CC FRAGMENT TYPE: internal
CC SEQUENCE 117 AA; 12389 MW; 76297 CN;

Query Match 58.5%; Score 48; DB 3; Length 117;
Best Local Similarity 50.0%; Pred. No. 2.18e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 22 GSYLPTAVELRY 33
:||||:|
1 ASYLSFSFLDY 12

RESULT 12
ID US-08-348-792-12 STANDARD: PRT; 329 AA.

AC xxxxxx
DT
XX Sequence 12, Application US/08348792
DE
XX Sequence 12, Application US/08348792
CC Patent No. 5576423
CC GENERAL INFORMATION:
CC APPLICANT: Aversa, Gregorio
CC APPLICANT: Chang, Chia-Chun J.
CC APPLICANT: Cocks, Benjamin G.
CC APPLICANT: de Vries, Jan E.
CC TITLE OF INVENTION: PURIFIED GENES ENCODING MAMMALIAN CELL
CC TITLE OF INVENTION: SURFACE ANTIGENS; PROTEINS AND ANTIBODIES
CC NUMBER OF SEQUENCES: 12
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: DNAX Research Institute
CC STREET: 901 California Avenue
CC CITY: Palo Alto
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94304-1104

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/348,792
CC FILING DATE: 02-DEC-1994
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Ching, Edwin P.
CC REGISTRATION NUMBER: 34,090
CC REFERENCE/DOCKET NUMBER: DX0436
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-852-9196
CC TELEFAX: 415-496-1200
CC INFORMATION FOR SEQ ID NO: 12:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 329 amino acids
CC TYPE: amino acid
CC MOLECULE TYPE: linear
CC TOPOLOGY: 1
CC SEQUENCE 329 AA; 36590 MW; 607424 CN;

Query Match 58.5%; Score 48; DB 1; Length 329;

Best Local Similarity 50.0%; Pred. No. 2.18e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 14 FLSLAFELSY 23
:||:|
3 YLSFSFLDY 12

RESULT 13
ID US-08-348-792-10 STANDARD: PRT; 343 AA.

AC xxxxxx
DT
XX

Sequence 10, Application US/08348792

CC Sequence 10, Application US/08348792
CC Patent No. 5576423
CC GENERAL INFORMATION:
CC APPLICANT: Aversa, Gregorio
CC APPLICANT: Chang, Chia-Chun J.
CC APPLICANT: Cocks, Benjamin G.
CC APPLICANT: de Vries, Jan E.
CC TITLE OF INVENTION: PURIFIED GENES ENCODING MAMMALIAN CELL
CC TITLE OF INVENTION: SURFACE ANTIGENS; PROTEINS AND ANTIBODIES
CC NUMBER OF SEQUENCES: 12
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: DNAX Research Institute
CC STREET: 901 California Avenue
CC CITY: Palo Alto
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94304-1104

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/348,792
CC FILING DATE: 02-DEC-1994
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Ching, Edwin P.
CC REGISTRATION NUMBER: 34,090
CC REFERENCE/DOCKET NUMBER: DX0436
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-852-9196
CC TELEFAX: 415-496-1200
CC INFORMATION FOR SEQ ID NO: 10:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 343 amino acids
CC TYPE: amino acid
CC MOLECULE TYPE: linear
CC TOPOLOGY: 1
CC SEQUENCE 343 AA; 38094 MW; 663822 CN;

Query Match 58.5%; Score 48; DB 1; Length 343;
Best Local Similarity 50.0%; Pred. No. 2.18e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 14 FLSLAFELSY 23
:||:|
3 YLSFSFLDY 12

RESULT 14
ID US-07-671-817A-4 STANDARD: PRT; 969 AA.

AC xxxxxx
DT

XX Sequence 4, Application US/07671817A
DE
XX Sequence 4, Application US/07671817A
CC Patent No. 5424409
CC GENERAL INFORMATION:
CC APPLICANT: Ely, Susan
CC APPLICANT: Tippet, Janet M
CC TITLE OF INVENTION: DNA constructs
CC NUMBER OF SEQUENCES: 6
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Cushman, Darby and Cushman
CC STREET: Eleventh floor, 1615 L Street, N.W.
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA
CC ZIP: 20036-3601
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/671,817A
CC FILING DATE: 19910401
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: GB 8823068.5
CC FILING DATE: 30-SEP-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/GB89/01157
CC FILING DATE: 29-SEP-1989
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202) 861-3000
CC TELEFAX: (202) 822-0944
CC TELEX: 6714627 CUSH
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 969 amino acids
CC TYPE: AMINO ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 969 AA; 109598 MW; 4754786 CN;
SQ
Query Match 58.5%; Score 48; DB 1; Length 969;
Best Local Similarity 60.0%; Pred. No. 2.18e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
653 TYLSDPFLD 662
: ||| | ||
OY 2 SYLSTSFSLD 11
RESULT 15
ID PCT-US95-05431-29 STANDARD: PRT: 1174 AA.
XX AC xxxxxx
XX
DT
XX
DE Sequence 29, Application PC/TUS9505431
CC
CC GENERAL INFORMATION:
CC APPLICANT:
CC APPLICANT: Street address: 5501 Oberlin Drive
CC APPLICANT: City: San Diego
CC APPLICANT: State/Province: California
CC APPLICANT: Country: US
CC APPLICANT: Postal code/zip: 92121
CC APPLICANT: Phone number: (619) 453-8030 Fax number: (619) 453-6991
CC APPLICANT: Telex number:

CC TITLE OF INVENTION: Improvement of Delta-Endotoxin Expression in
CC TITLE OF INVENTION: Pseudomonas fluorescens
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: David R. Salivanchik
CC STREET: 2421 N.W. 41st Street, Suite A-1
CC CITY: Gainesville
CC STATE: Florida
CC COUNTRY: USA
CC ZIP: 32606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/05431
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Salivanchik, David R.
CC REGISTRATION NUMBER: 31,794
CC REFERENCE/DOCKET NUMBER: MA83
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (904) 375-8100
CC TELEFAX: (904) 372-5800
CC INFORMATION FOR SEQ ID NO: 29:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1174 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1174 AA; 133119 MW; 7052059 CN;
SQ
Query Match 58.5%; Score 48; DB 3; Length 1174;
Best Local Similarity 60.0%; Pred. No. 2.18e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
DB 646 TYLSDPFLD 655
: ||| | ||
OY 2 SYLSTSFSLD 11

Search completed: Thu Sep 2 12:40:46 1999
Job time : 7 secs.



CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 12 AA;

Query Match 100.0%; Score 78; DB 27; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.96e-01;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 asy1stss1hy 12
 |||||
 QY 1 ASY1STSSSLHY 12

RESULT 2
 W27588 standard; peptide: 12 AA.

AC W27588:
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.

OS Homo sapiens.
 PN WO929131-A1.
 PD 14-AUG-1997; 002219.
 PF 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRM, Kaymakcalan Z, Labkovsky B,
 PI Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR WPI: 97-415302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 Claim 20; Page 73; 102pp; English.
 The present sequence is a novel anti-human tumour necrosis
 factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 12 AA;

Query Match 88.5%; Score 69; DB 27; Length 12;
 Best Local Similarity 91.7%; Pred. No. 1.85e-00;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 1 asy1stss1dy 12
 |||||
 QY 1 ASY1STSSSLHY 12

RESULT 3
 W27593 standard; peptide: 12 AA.

AC W27593:
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.

OS Homo sapiens.
 PN WO929131-A1.
 PD 14-AUG-1997; 002219.
 PF 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRM, Kaymakcalan Z, Labkovsky B,
 PI Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR WPI: 97-415302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 Claim 20; Page 75; 102pp; English.
 The present sequence is a novel anti-human tumour necrosis
 factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 12 AA;

Query Match 85.9%; Score 67; DB 27; Length 12;
 Best Local Similarity 83.3%; Pred. No. 3.03e+00;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1 asy1stasse1y 12
 |||||
 QY 1 ASY1STSSSLHY 12

RESULT 4
 W27592 standard; peptide: 12 AA.

AC W27592:
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;

KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997: U02219.
 PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuiness BR, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WP: 97-415302/38.
 CC High affinity antibodies against human TNF alpha - useful to inhibit
 CC TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20: Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA;
 SQ

Query Match 82.1%; Score 64; DB 27; Length 12;
 Best Local Similarity 83.3%; Pred. No. 6.27e+00;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 asfistseslay 12
 ||:|||||||
 QY 1 ASYSTSSSLHY 12

RESULT 5
 ID W27569 standard; Protein; 121 AA.
 AC W27569;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain variable region.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody;
 KW heavy chain; variable region; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997: U02219.
 PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuiness BR, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WP: 97-415302/38.
 CC High affinity antibodies against human TNF alpha - useful to inhibit
 CC TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20: Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).

PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuiness BR, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WP: 97-415302/38.
 DR N-PSDB; T88404.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 16: Page 76; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain variable region.
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 121 AA;
 SQ

Query Match 80.8%; Score 63; DB 27; Length 121;
 Best Local Similarity 81.8%; Pred. No. 7.99e+00;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Dh 100 sylvstasldy 110
 |||||:||||
 QY 2 SYLSTSSSLHY 12

RESULT 6
 ID W27590 standard; peptide; 12 AA.
 AC W27590;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997: U02219.
 PR 25-NOV-1996: US-031476.
 PR 09-FEB-1996: US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankovich JA, McGuiness BR, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WP: 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20: Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA.

Query Match 78.2%; Score 61; DB 27; Length 12;
Best Local Similarity 83.3%; Pred. No. 1.29e+01;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DB 1 asy1stsfldy 12
Oy 1 ASYLSTSSSLMT 12

RESULT 7
ID W27586 standard; peptide: 12 AA.
AC W27586;
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
PD 14-AUG-1997.
PF 10-FEB-1997; U02219.
PR 25-NOV-1996; US-031476.
PS Claim 20; Page 73; 102pp; English.
PT The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC

CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA.

Query Match 73.1%; Score 57; DB 27; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.33e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 asy1stsssl 10
Oy 1 ASYLSTSSSL 10

RESULT 8
ID W27589 standard; peptide: 12 AA.
AC W27589;
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
PD 14-AUG-1997.
PF 10-FEB-1997; U02219.
PR 25-NOV-1996; US-031476.
PS Claim 20; Page 73; 102pp; English.
PT The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA.
Query Match 73.1%; Score 57; DB 27; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.33e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 asy1stsssl 10
Oy 1 ASYLSTSSSL 10

RESULT 9
ID W27587 standard; peptide; 12 AA.
AC W27587;
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HIVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
14-AUG-1997.
10-FEB-1997: U02219.
25-NOV-1996: US-031476.
09-FEB-1996: US-599226.
PR (BADT) BASE AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 73: 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L2929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption diseases,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
CC Sequence 12 AA:
SQ

Query Match 73.1%; Score 57; DB 27; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.33e+01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 asy1stss1 10
OY 1 ASYSTSSSL 10

RESULT 10
ID W34987 standard; Protein; 875 AA.
AC W34987;
DT 21-MAY-1998 (first entry)
DE Bankia gouldi endoglucanase.
KW Endoglucanase; cellulase; carboxymethylcellulose; cellulose;
KW biomass; beta-1,4-glycosidic bond; hydrolysis; saccharification;
KW thermostable enzyme; thermophilic; glycosidase.
OS Bankia gouldi mix (Clone 37Gp2).
PN W09744361-A1.
27-NOV-1997.
PD

PF 22-MAY-1997; U08793.
PA 22-MAY-1996; US-651572.
PR (RECO-) RECOMBINANT BIOCATALYSIS INC.
PI Lam DE, Mathur EJ;
DR WPI: 98-018435/02.
DR N-PSDB: T94195.
PT Endoglucanase(s), preferably from archaeal bacterium, ABPII 1a -
PT useful to degrade carboxymethylcellulose and hydrolyse of
PT beta-1,4-glycosidic bonds in cellulose
PS Claim 1: Fig 1c: 164pp: English.
CC This protein comprises an endoglucanase of Bankia gouldi mix (clone
CC 37Gp2) that is capable of degrading carboxymethylcellulose and of
CC hydrolysing the beta-1,4-glycosidic bonds in cellulose. It has
CC homology to an endoglucanase of archaeobacterium AEPiIIa (see
CC W34985). It can be produced from native cells or from recombinant
CC host cells, especially prokaryotic host cells transformed with a
CC plasmid or virus-derived vector including the endoglucanase DNA
CC (see T94195). 24 Endoglucanases (see W34986-W35008) are claimed.
CC They can be used to degrade cellulose for the conversion of plant
CC biomass into fuels and chemicals, for use in detergents, textiles,
CC animal feed, waste treatment, and in the fruit juice and brewing
CC industries for the clarification and extraction of juices.
CC Sequence 875 AA:
SQ

Query Match 66.7%; Score 52; DB 28; Length 875;
Best Local Similarity 58.3%; Pred. No. 1.06e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 618 psy1tvds1 629
OY 1 ASYSTSSSLHY 12

RESULT 11
ID W27594 standard; peptide; 12 AA.
AC W27594;
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HIVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
14-AUG-1997.
10-FEB-1997: U02219.
25-NOV-1996: US-031476.
09-FEB-1996: US-599226.
PR (BADT) BASE AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
PS Disclosure: Page 75: 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L2929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple

CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 12 AA;

Query Match 65.4%; Score 51; DB 27; Length 12;
 Best Local Similarity 88.9%; Pred. No. 1.33e+02;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 2 sylstas1 10
 |||||:
 QY 2 SYLSTSSSL 10

RESULT 12
 ID W27563 standard; peptide; 12 AA.

AC W27563:
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 KS
 FT Key Location/Qualifiers
 FT MISC.difference 12
 FT /label= Tyr, Asn
 FT WC09729131-A1.
 PD 14-AUG-1997.
 PE 10-FEB-1997; U02219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Markovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 PI WPI: 97-413302/38.
 FT High affinity antibodies against human TNF alpha - useful to inhibit
 FT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 9; Page 65; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 12 AA;

Query Match 65.4%; Score 51; DB 27; Length 12;
 Best Local Similarity 88.9%; Pred. No. 1.33e+02;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 2 sylstas1 10
 |||||:
 QY 2 SYLSTSSSL 10

RESULT 13

ID W21664 standard; Protein; 321 AA.

AC W21664;

DT 29-SEP-1997 (first entry)

DE Rat spermatic chemoreceptor D-8.

KW Sperm receptor; spermatic chemoreceptor; contraceptive; vaccine;

KW infertility.

OS Rattus sp.

FT Key Location/Qualifiers

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

Query Match 65.4%; Score 51; DB 27; Length 12;
 Best Local Similarity 88.9%; Pred. No. 1.33e+02;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 2 sylstas1 10
 |||||:
 QY 2 SYLSTSSSL 10

RESULT 13

ID W21664 standard; Protein; 321 AA.

AC W21664;

DT 29-SEP-1997 (first entry)

DE Rat spermatic chemoreceptor D-8.

KW Sperm receptor; spermatic chemoreceptor; contraceptive; vaccine;

KW infertility.

OS Rattus sp.

FT Key Location/Qualifiers

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

THIS PAGE BLANK (USPTO)

MUSE

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch.p protein - protein database search, using Smith-Waterman algorithm
on: Thu Sep 2 12:44:21 1999: Maspar time 1.37 Seconds
88.617 Million cell updates/sec
Abular output not generated.

Title: >US-08-599-226-32
Description: (1-12) from US08599226.Pep
Perfect Score: 78
Sequence: 1 ASYLSTSSSLHY 12

Scoring table: PAM 150
Gap 15

Searched: 106580 seqs, 10152877 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1-5A_COMB 2:5B_COMB 3:PCF9_COMB 4:backfiles1

Statistics: Mean 16.037; Variance 54.698; scale 0.293

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	49	62.8	400	2	US-08-103-Sequence 9, Applicatio	9.84e+01
2	49	62.8	491	1	US-08-462-Sequence 2, Applicatio	9.84e+01
3	49	62.8	491	1	US-08-176-Sequence 2, Applicatio	9.84e+01
4	48	61.5	29	1	US-08-066-Sequence 59, Applicati	1.23e+02
5	48	61.5	158	2	US-08-828-Sequence 3, Applicatio	1.23e+02
6	48	61.5	159	2	US-08-828-Sequence 1, Applicatio	1.23e+02
7	48	61.5	362	3	PCT-US93-0Sequence 16, Applicati	1.23e+02
8	48	61.5	362	1	US-08-118-Sequence 2, Applicatio	1.23e+02
9	48	61.5	388	1	US-08-087-Sequence 2, Applicatio	1.23e+02
10	48	61.5	400	1	US-07-916-Sequence 6, Applicatio	1.23e+02
11	48	61.5	400	1	US-08-351-Sequence 4, Applicatio	1.23e+02
12	48	61.5	400	1	US-08-351-Sequence 5, Applicatio	1.23e+02
13	48	61.5	400	1	US-07-783-Sequence 1, Applicatio	1.23e+02
14	48	61.5	402	1	US-08-444-Sequence 6, Applicatio	1.23e+02
15	48	61.5	402	1	US-08-087-Sequence 15, Applicati	1.23e+02
16	48	61.5	405	1	US-08-351-Sequence 2, Applicatio	1.23e+02
17	48	61.5	408	1	US-07-916-Sequence 2, Applicatio	1.23e+02
18	48	61.5	408	1	US-08-351-Sequence 3, Applicatio	1.23e+02
19	48	61.5	1046	2	US-08-386-Sequence 2, Applicatio	1.23e+02
20	48	61.5	1255	2	US-08-414-Sequence 2, Applicatio	1.23e+02
21	48	61.5	1255	2	US-08-625-Sequence 2, Applicatio	1.23e+02
22	48	61.5	1255	1	US-08-467-Sequence 68, Applicati	1.23e+02
23	48	61.5	1255	2	US-08-468-Sequence 68, Applicati	1.23e+02

RESULT	1	ALIGNMENTS
ID	US-08-103-170-9	STANDARD: PRT: 400 AA.
XX	xxxxxx	
XX		
DT		
DE	Sequence 9, Application US/08103170	
CC	Sequence 9, Application US/08103170	
CC	Patent No. 5885824	
CC	GENERAL INFORMATION:	
CC	APPLICANT: Yamada, Tadataka	
CC	APPLICANT: Gantuz, Ira	
CC	TITLE OF INVENTION: Recombinant Genomic Clones Encoding	
CC	TITLE OF INVENTION: Histamine H1, H2, and H3 Receptors, Methods For Product	
CC	TITLE OF INVENTION: Thereof, and Proteins Encoded Therefrom	
CC	NUMBER OF SEQUENCES: 41	
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,	
CC	ADDRESSEE: P.C.	
CC	STREET: 1755 Jefferson Davis Highway, Fourth Floor	
CC	CITY: Arlington	
CC	STATE: Virginia	
CC	COUNTRY: U.S.A.	
CC	ZIP: 22202	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	
CC	COMPUTER: IBM PC compatible	
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	
CC	SOFTWARE: Patent Release #1.0, Version #1.25	
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER: US/08/103,170	
CC	FILING DATE:	
CC	CLASSIFICATION: 435	
CC	PRIOR APPLICATION DATA:	
CC	APPLICATION NUMBER: US 07/633,060	
CC	FILING DATE: 24-DEC-1990	
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME: Lavelle, Jean-Paul	
CC	REGISTRATION NUMBER: 31,451	
CC	REFERENCE/DOCKET NUMBER: 2363-017-55	
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE: (703)521-4500	
CC	TELEFAX: (703)486-2347	
CC	TELEX: 24885 OPAT UR	
CC	INFORMATION FOR SEQ ID NO: 9:	

CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 400 amino acids
CC TYPE: amino acid
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC ORIGINAL SOURCE:
CC ORGANISM: Homo sapiens
SQ SEQUENCE 400 AA; 42804 MW; 785909 CN;

Query Match 62.8%; Score 49; DB 1; Length 400;
Best Local Similarity 40.0%; Pred. No. 9.84e+01;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 134 YLAVTNTLRY 143
||:|:|:|:|
QY 3 YLSTSSSLHY 12

RESULT 2
US-08-462-733-2 STANDARD; PRT: 491 AA.

AC xxxxxx

Sequence 2, Application US/08462733

CC Patent No. 5610019

CC GENERAL INFORMATION:

CC APPLICANT: Day, Joseph R.

CC APPLICANT: Albers, John J.

CC APPLICANT: Lofton-Day, Catherine E.

CC APPLICANT: Adolphson, Janet L.

CC TITLE OF INVENTION: Phospholipid Transfer Proteins

CC NUMBER OF SEQUENCES: 9

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Zymogenetics, Inc.

CC STREET: 4225 Roosevelt Way, N.E.

CC CITY: Seattle

CC STATE: WA

CC COUNTRY: USA

CC ZIP: 98105

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/462,733

CC FILING DATE:

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Parker, Gary E

CC REGISTRATION NUMBER: 31-648

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 206-547-8080 ext 322

CC TELEFAX: 206-548-2329

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 491 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

SQ SEQUENCE 491 AA; 54569 MW; 1262097 CN;

Query Match 62.8%; Score 49; DB 1; Length 491;
Best Local Similarity 33.3%; Pred. No. 9.84e+01;
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Db 453 AGFLITGADLHF 464
|:|:|:|:|:|

QY 1 ASYLSTSSSLHY 12

RESULT 3
ID US-08-176-402-2 STANDARD; PRT: 491 AA.

AC xxxxxx

Sequence 2, Application US/08176402

CC Patent No. 5622843

CC GENERAL INFORMATION:

CC APPLICANT: Day, Joseph R.

CC APPLICANT: Albers, John J.

CC APPLICANT: Lofton-Day, Catherine E.

CC TITLE OF INVENTION: Phospholipid Transfer Proteins

CC NUMBER OF SEQUENCES: 9

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Zymogenetics, Inc.

CC STREET: 4225 Roosevelt Way, N.E.

CC CITY: Seattle

CC STATE: WA

CC COUNTRY: USA

CC ZIP: 98105

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/176,402

CC FILING DATE:

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Parker, Gary E

CC REGISTRATION NUMBER: 31-684

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 206-547-8080 ext 322

CC TELEFAX: 206-548-2329

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 491 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

SQ SEQUENCE 491 AA; 54569 MW; 1262097 CN;

Query Match 62.8%; Score 49; DB 1; Length 491;
Best Local Similarity 33.3%; Pred. No. 9.84e+01;
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Db 453 AGFLITGADLHF 464
QY 1 ASYLSTSSSLHY 12

RESULT 4
ID US-08-066-325-59 STANDARD; PRT: 29 AA.

AC xxxxxx

Sequence 59, Application US/08066325

CC Patent No. 5667967

CC GENERAL INFORMATION:

CC APPLICANT: Steinman, Lawrence
CC APPLICANT: Oksenberg, Jorge
CC APPLICANT: Bernard, Claude
CC TITLE OF INVENTION: T-CELL RECEPTOR VARIABLE TRANSCRIPTS AS DISEASE RELATED MA
CC NUMBER OF SEQUENCES: 157
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SEED and BERRY LLP
CC STREET: 6300 Columbia Center, 701 Fifth Avenue
CC CITY: Seattle
CC STATE: Washington
CC COUNTRY: USA
CC ZIP: 98104-7092
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/066,325
CC FILING DATE: 21-MAY-1993
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: No. 5667967tenburg Ph.D., Carol
CC REGISTRATION NUMBER: 39,317
CC REFERENCE/DOCKET NUMBER: 690068.408C1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (206) 622-4900
CC TELEFAX: (206) 682-6031
CC INFORMATION FOR SEQ ID NO: 59:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 29 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 29 AA; 3155 MW; 4596 CN;
SQ

Query Match 61.5%; Score 48; DB 1; Length 29;
Best Local Similarity 41.7%; Pred. No. 1.23e+02;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 4 SSFLGNSPLHF 15
:|:|:|:|:|:
QY 1 ASYLSTSSSLHY 12

RESULT 5
US-08-828-832-3 STANDARD: PRT: 158 AA.
xxxxxx

Sequence 3, Application US/08828832
Sequence 3, Application US/08828832
Patent No. 5827711
GENERAL INFORMATION:
CC APPLICANT: Lal, Preeti
CC APPLICANT: Shah, Purvi
CC TITLE OF INVENTION: NOVEL SUCCINATE DEHYDROGENASE SUBUNIT
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS

CC SOFTWARE: FastSeq for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/828,832
CC FILING DATE: Herewith
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: PF-0250 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 158 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: GenBank
CC CLONE: 1575011
SQ SEQUENCE 158 AA; 17096 MW; 135076 CN;

Query Match 61.5%; Score 48; DB 2; Length 158;
Best Local Similarity 41.7%; Pred. No. 1.23e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 81 AAYLPCSAMDY 92
|:|:|:|:|:
QY 1 ASYLSTSSSLHY 12

RESULT 6
US-08-828-832-1 STANDARD: PRT: 159 AA.
xxxxxx

Sequence 1, Application US/08828832
Sequence 1, Application US/08828832
Patent No. 5827711
GENERAL INFORMATION:
CC APPLICANT: Lal, Preeti
CC APPLICANT: Shah, Purvi
CC TITLE OF INVENTION: NOVEL SUCCINATE DEHYDROGENASE SUBUNIT
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/828,832
CC FILING DATE: Herewith
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.

CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: PF-0250 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 159 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: Consensus
CC CLONE: 2454416
CC
SQ SEQUENCE 159 AA: 17043 MW: 129477 CN;

Query Match 61.5%; Score 48; DB 2; Length 159;
Best Local Similarity 41.7%; Pred. No. 1.23e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 82 AAVLNPCSAMDY 93
|:|:|:|:|:
QY 1 ASYLTSSSLHY 12

RESULT 7
XX PCT-US93-08528-16 STANDARD; PRT; 362 AA.
XX xxxxxx

Sequence 16, Application PC/TUS9308528
GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

CC MOLECULE TYPE: peptide
SQ SEQUENCE 362 AA: 38722 MW: 635563 CN;

Query Match 61.5%; Score 48; DB 3; Length 362;
Best Local Similarity 40.0%; Pred. No. 1.23e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 98 YLAVTNPLRY 107
|:|:|:|:|:
QY 3 YLTSSSLHY 12

RESULT 8
XX US-08-118-270-16 STANDARD; PRT; 362 AA.
XX xxxxxx

Sequence 16, Application US/08118270
GENERAL INFORMATION:
PATENT NO. 5508384
APPLICANT: Murphy, Randall B.
APPLICANT: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE 362 AA: 38722 MW: 635563 CN;

Query Match 61.5%; Score 48; DB 1; Length 362;
Best Local Similarity 40.0%; Pred. No. 1.23e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 98 YLAVTNPLRY 107
|:|:|:|:|:
QY 3 YLTSSSLHY 12

RESULT 9
ID US-08-087-772A-2 STANDARD; PRT: 388 AA.
XX
AC xxxxxx
XX
XX
DE Sequence 2, Application US/08087772A
XX
CC Sequence 2, Application US/08087772A
CC Patent No. 5651155
CC GENERAL INFORMATION:
CC APPLICANT: Namimas, Clara
CC APPLICANT: Emorine, Jean L.
CC APPLICANT: Strosberg, Donny A.
CC TITLE OF INVENTION: Nucleotide Sequences Encoding the Murine
CC TITLE OF INVENTION: Beta3-Adrenergic Receptor and Their Applications
CC NUMBER OF SEQUENCES: 17
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Bell, Seltzer, Park & Gibson
CC STREET: Post Office Drawer 34009
CC CITY: Charlotte
CC STATE: No. 5691155th Carolina
CC COUNTRY: USA
CC ZIP: 28234
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/087,772A
CC FILING DATE:
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Linker, Raymond O.
CC REGISTRATION NUMBER: 26,419
CC REFERENCE/DOCKET NUMBER: 3339-195
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 919-881-3140
CC TELEFAX: 919-881-3175
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 388 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 388 AA; 41667 MW; 771830 CN;
Query Match 61.5%; Score 48; DB 1; Length 388;
Best Local Similarity 40.0%; Pred. No. 1.23e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Db 133 YLAVTNPLRY 142
||:|:|:|:|
QY 3 YLSTSSSLHY 12
RESULT 10
ID US-07-916-901-6 STANDARD; PRT: 400 AA.
XX
AC xxxxxx
XX
XX
DE Sequence 6, Application US/07916901
XX
CC Sequence 6, Application US/07916901
CC Patent No. 5364772
CC GENERAL INFORMATION:
CC APPLICANT: Granneman, James G.
CC APPLICANT: Lahners, Kristine N.
CC APPLICANT: Rao, Donald D.

CC TITLE OF INVENTION: @ @3-ADRENERGIC RECEPTOR PROTEIN AND DNA
CC TITLE OF INVENTION: ENCODING SAME
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: REISING, ETHINGTON, BARNARD, PERRY &
CC ADDRESSEE: MILTON
CC STREET: 201 W. Big Beaver - Ste. 400; P.O. Box 4390
CC CITY: Troy
CC STATE: Michigan
CC COUNTRY: USA
CC ZIP: 48099
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/916,901
CC FILING DATE: 19920720
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Kohn, Kenneth I.
CC REGISTRATION NUMBER: 30,955
CC REFERENCE/DOCKET NUMBER: P-324(WSU)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (313) 689-3554
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 400 amino acids
CC TYPE: AMINO ACID
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 400 AA; 43146 MW; 840185 CN;
Query Match 61.5%; Score 48; DB 1; Length 400;
Best Local Similarity 40.0%; Pred. No. 1.23e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Db 133 YLAVTNPLRY 142
||:|:|:|:|
QY 3 YLSTSSSLHY 12
RESULT 11
ID US-08-351-473B-4 STANDARD; PRT: 400 AA.
XX
AC xxxxxx
XX
XX
DE Sequence 4, Application US/08351473B
XX
CC Sequence 4, Application US/08351473B
CC Patent No. 5656440
CC GENERAL INFORMATION:
CC APPLICANT: LENZEN, GERLINDA
CC APPLICANT: KAPOOR, ARCHANA
CC TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE
CC TITLE OF INVENTION: BOVINE BETA3-ADRENERGIC RECEPTOR AND THEIR APPLICATIONS
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: OBLOIN, SPIYAK, MCLELLAND, MAIER & NEUSTADT
CC STREET: 1735 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CC CITY: ARLINGTON
CC STATE: VIRGINIA
CC COUNTRY: USA
CC ZIP: 22202
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/351.473B
CC FILING DATE: 21-FEB-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 93 04670
CC FILING DATE: 21-APR-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/FR94/00447
CC FILING DATE: 21-APR-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: OBLON, NORMAN F.
CC REGISTRATION NUMBER: 24,618
CC REFERENCE/DOCKET NUMBER: 6639-001-0X PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703) 413-3000
CC TELEFAX: (703) 413-2220
CC TELEX: 248855 OPAT UR
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 400 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 400 AA; 4316 MW; 839778 CN;
SQ

Query Match 61.5%; Score 48; DB 1; Length 400;
Best Local Similarity 40.0%; Pred. No. 1.23e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 133 YIAVNP142
||:..:|:|
OY 3 YLSTSSSLHY 12

RESULT 12
ID US-08-351-473B-5 STANDARD: PRT; 400 AA.
AC xxxxxx
XX
DT
XX
DE Sequence 5, Application US/08351473B
XX Sequence 5, Application US/08351473B
CC Patent No. 5656440
CC GENERAL INFORMATION:
CC APPLICANT: LENZEN, GERLINDA
CC APPLICANT: KAPOOR, ARCHANA
CC TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE
CC TITLE OF INVENTION: BOVINE BETAS-ADRENERGIC RECEPTOR AND THEIR APPLICATIONS
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
CC STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CC CITY: ARLINGTON
CC STATE: VIRGINIA
CC COUNTRY: USA
CC ZIP: 22202
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/351.473B
CC FILING DATE: 21-FEB-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 93 04670
CC FILING DATE: 21-APR-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/FR94/00447

CC FILING DATE: 21-APR-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: OBLON, NORMAN F.
CC REGISTRATION NUMBER: 24,618
CC REFERENCE/DOCKET NUMBER: 6639-001-0X PCT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703) 413-3000
CC TELEFAX: (703) 413-2220
CC TELEX: 248855 OPAT UR
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 400 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 400 AA; 43005 MW; 820814 CN;
SQ

Query Match 61.5%; Score 48; DB 1; Length 400;
Best Local Similarity 40.0%; Pred. No. 1.23e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 133 YIAVNP142
||:..:|:|
OY 3 YLSTSSSLHY 12

RESULT 13
ID US-07-783-602C-1 STANDARD: PRT; 400 AA.
AC xxxxxx
XX
DT
XX
DE Sequence 1, Application US/07783602C
XX Sequence 1, Application US/07783602C
CC Patent No. 5418160
CC GENERAL INFORMATION:
CC APPLICANT: J. Craig Venter et al
CC TITLE OF INVENTION: A FAT CELL SPECIFIC a-ADRENERGIC
CC TITLE OF INVENTION: RECEPTOR
CC NUMBER OF SEQUENCES: 1
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Lowe, Price, LeBlanc & Becker
CC STREET: Suite 300, 99 Canal Center Plaza
CC CITY: Alexandria
CC STATE: Virginia
CC COUNTRY: USA
CC ZIP: 22314
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: DOS Text File
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/783,602C
CC FILING DATE: 19911101
CC CLASSIFICATION: 436
CC ATTORNEY/AGENT INFORMATION:
CC NAME: J.G. Mullins
CC REGISTRATION NUMBER: 33073
CC REFERENCE/DOCKET NUMBER: 717-098
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 703 684 1111
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 400
CC TYPE: AMINO ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: Polypeptide
CC SEQUENCE 400 AA; 4316 MW; 843002 CN;
SQ

Query Match 61.5%; Score 48; DB 1; Length 400;
Best Local Similarity 40.0%; Pred. No. 1.23e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

DB 133 YLAVTNPLRY 142
||: :||:|
QY 3 YLSTSSSLHY 12

RESULT 14
ID US-08-444-734A-6 STANDARD; PRT: 402 AA.
AC xxxxxx
XX
XX
XX
XX
XX

Sequence 6, Application US/08444734A
Patent No. 5610282
GENERAL INFORMATION:
APPLICANT: Sibley, David R.
APPLICANT: Monsma, Frederick J.
APPLICANT: Mahan, Lawrence C.
APPLICANT: McVittie, Loris D.
TITLE OF INVENTION: cDNA encoding the rat D1 dopamine
TITLE OF INVENTION: receptor linked to adenylyl cyclase activation and
TITLE OF INVENTION: expression of the receptor protein in plasmid-transfected
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,734A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/029,917
FILING DATE: 03-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/548,714
FILING DATE: 06-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: NIH065.001FW1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (714) 760-0404
TELEFAX: (714) 760-9502
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
SEQUENCE 402 AA; 42931 MW; 794255 CN;

Query Match 61.5%; Score 48; DB 1; Length 402;

Best Local Similarity 40.0%; Pred. No. 1.23e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

DB 136 YLAVTNPLRY 145
||: :||:|
QY 3 YLSTSSSLHY 12

RESULT 15
ID US-08-087-772A-15 STANDARD; PRT: 402 AA.
AC xxxxxx
XX
XX
XX
XX
XX

Sequence 15, Application US/08087772A
Patent No. 5691155
GENERAL INFORMATION:
APPLICANT: Nahmias, Clara
APPLICANT: Emorine, Jean L.
APPLICANT: Strosberg, Donny A.
TITLE OF INVENTION: Nucleotide Sequences Encoding the Murine
TITLE OF INVENTION: Beta3-Adrenergic Receptor and Their Applications
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Bell, Seltzer, Park & Gibson
STREET: Post Office Drawer 34009
CITY: Charlotte
STATE: NC 28234
COUNTRY: USA
ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,772A
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Linker, Raymond O.
REGISTRATION NUMBER: 26,419
REFERENCE/DOCKET NUMBER: 3339-195
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-881-3140
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE 402 AA; 42931 MW; 794255 CN;

Query Match 61.5%; Score 48; DB 1; Length 402;
Best Local Similarity 40.0%; Pred. No. 1.23e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

DB 136 YLAVTNPLRY 145
||: :||:|
QY 3 YLSTSSSLHY 12

Search completed: Thu Sep 2 12:44:28 1999
Job time : 7 secs.

THIS PAGE BLANK (USPTO)

 WIRESEED (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch.p protein - protein database search, using Smith-Waterman algorithm
 on: Thu Sep 2 12:43:10 1999; MasPar time 2.22 Seconds
 153.061 Million cell updates/sec
 Similar output not generated.

Title: >US-08-599-226-32
 Description: (1-12) from US08599226.pep
 Perfect Score: 78
 Sequence: 1 ASYLSTSSSLHY 12

Scoring table: PAM 150
 Gap 15

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot37
 1:swissprot

Statistics: Mean 24.890; Variance 28.198; scale 0.883

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	52	66.7	594	1	COX1_PHYPO CYTOCHROME C OXIDASE P	2.45e+00
2	52	66.7	828	1	YKR6_YEAST HYPOTHETICAL 93.4 KD P	2.45e+00
3	51	65.4	303	1	YNA6_YEAST HYPOTHETICAL 34.0 KD T	4.01e+00
4	51	65.4	428	1	B4AR_MELGA BETA-4C ADRENERGIC REC	4.01e+00
5	50	65.4	543	1	YJF8_YEAST HYPOTHETICAL 50.8 KD P	4.01e+00
6	50	64.1	328	1	YGJR_ECOLI HYPOTHETICAL 36.2 KD P	6.49e+00
7	50	64.1	445	1	HPPD_ARATH 4-HYDROXYPHENYLPYRUVAT	6.49e+00
8	50	64.1	1010	1	CLPP_CLEU PROBABLE CLPP-LIKE PRO	6.49e+00
9	49	62.8	493	1	PLTP_HUMAN PHOSPHOLIPID TRANSFER	1.04e+01
10	49	62.8	2410	1	POL1_BAYM GENOME POLYPROTEIN 1	1.04e+01
11	49	62.8	2412	1	POL1_BAYM GENOME POLYPROTEIN 1	1.04e+01
12	48	61.5	158	1	DHSD_BOVIN SUCCINATE DEHYDROGENAS	1.67e+01
13	48	61.5	159	1	DHSD_HUMAN SUCCINATE DEHYDROGENAS	1.67e+01
14	48	61.5	351	1	B3AR_CAVPO BETA-3 ADRENERGIC REC	1.67e+01
15	48	61.5	400	1	B3AR_MOUSE BETA-3 ADRENERGIC REC	1.67e+01
16	48	61.5	400	1	B3AR_MOUSE BETA-3 ADRENERGIC REC	1.67e+01
17	48	61.5	405	1	B3AR_BOVIN BETA-3 ADRENERGIC REC	1.67e+01
18	48	61.5	405	1	B3AR_BOVIN BETA-3 ADRENERGIC REC	1.67e+01
19	48	61.5	408	1	B3AR_HUMAN BETA-3 ADRENERGIC REC	1.67e+01
20	48	61.5	418	1	B3AR_MACMU BETA-3 ADRENERGIC REC	1.67e+01
21	48	61.5	441	1	SECV_MYCTU PREPROTEIN TRANSLOCASE	1.67e+01
22	48	61.5	590	1	CTP SYNTHASE (EC 6.3.4	1.67e+01
23	48	61.5	963	1	IRE2_RAT IRON-RESPONSIVE ELEMENT	1.67e+01

24	48	61.5	963	1	IRE2_HUMAN IRON-RESPONSIVE ELEMENT	1.67e+01
25	48	61.5	1038	1	SOG_DROME DORSAL-VENTRAL PATTERN	1.67e+01
26	48	61.5	1103	1	RETINAL_GNATYLYL_CYCIA	1.67e+01
27	48	61.5	1255	1	ERR2_HUMAN ERBB-2 RECEPTOR PROTEI	1.67e+01
28	48	61.5	1469	1	DP77_CAEBL CHROMOSOME CONDENSATIO	1.67e+01
29	47	60.3	112	1	Y143_MYGE HYPOTHETICAL PROTEIN M	2.64e+01
30	47	60.3	291	1	HMP1_BOVIN PITUITARY-SPECIFIC POS	2.64e+01
31	47	60.3	291	1	HMP1_MOUSE PITUITARY-SPECIFIC POS	2.64e+01
32	47	60.3	291	1	HMP1_HUMAN PITUITARY-SPECIFIC POS	2.64e+01
33	47	60.3	291	1	HMP1_PIG PITUITARY-SPECIFIC POS	2.64e+01
34	47	60.3	291	1	HMP1_SHEEP PITUITARY-SPECIFIC POS	2.64e+01
35	47	60.3	394	1	BENE_ACICA BENOATE MEMBRANE TRAN	2.64e+01
36	47	60.3	434	1	KE51_YEAST KES1 PROTEIN	2.64e+01
37	47	60.3	533	1	YADC_SCHPO HYPOTHETICAL 62.2 KD P	2.64e+01
38	47	60.3	646	1	COAT_ASPB7 MAJOR CAPSID PROTEIN (2.64e+01
39	47	60.3	684	1	Y118_MYCTU HYPOTHETICAL 57.3 KD P	2.64e+01
40	47	60.3	1941	1	YRM8_CAEBL HYPOTHETICAL 216.3 KD	2.64e+01
41	46	59.0	132	1	YMO7_MYCTU HYPOTHETICAL 14.7 KD P	4.14e+01
42	46	59.0	184	1	YBET_ECOLI HYPOTHETICAL 20.9 KD P	4.14e+01
43	46	59.0	435	1	Y113_CAEBL HYPOTHETICAL 48.6 KD P	4.14e+01
44	46	59.0	530	1	YAN2_SCHPO HYPOTHETICAL 54.2 KD S	4.14e+01
45	46	59.0	1844	1	POLR_TYMYC RNA REPLICASE POLYPROT	4.14e+01

ALIGNMENTS

RESULT ID	1	STANDARD:	PRT:	594 AA.
AC	007434;			
DT	01-FEB-1996 (REL. 33, CREATED)			
DT	01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)			
DT	01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)			
DE	CYTOCHROME C OXIDASE POLYPEPTIDE I (EC 1.9.3.1).			
GN	COXI.			
OS	PHYSARUM POLYCEPHALUM (SLIME MOLD).			
OC	MITOCHONDRION.			
OG	EUKARYOTA; MYXOMYCETES; PHYSARIDA; PHYSARUM.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MC;			
RA	MEDLINE; 94064614.			
RX	GOTT J.M., VISOMIRSKI L.M., HUNTER J.L.;			
RT	"Substitutional and insertional RNA editing of the cytochrome c			
RT	oxidase subunit 1 mRNA of Physarum polycephalum."			
RL	J. BIOL. CHEM. 268:25483-25486(1993).			
CC	- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY			
CC	CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNIT 1-			
CC	3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE			
CC	CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN			
CC	CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2			
CC	AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3			
CC	AND COPPER B.			
CC	- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +			
CC	4 FERROCYTOCHROME C.			
CC	- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.			
CC	- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL			
CC	INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS.			
CC	- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.			
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
DR	EMBL; L14769; E101261; -			
DR	PROSITE; PS00077; COX1; 1.			
DR	PFAM; PF00115; COX1; 2.			
DR	HSSP; P00396; 10CC.			
KW	OXIDOREDUCTASE; HEME; COPPER; MITOCHONDRION; TRANSMEMBRANE;			
KW	RESPIRATORY CHAIN; INNER MEMBRANE.			

RL J. BIOL. CHEM. 269:24810-24819(1994).
 CC -1- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDATE THE CATECHOLAMINE-
 CC INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
 CC PROTEINS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: BROAD TISSUE DISTRIBUTION.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U13977; G556604; -
 CC EMBL: U13978; G555882; -
 CC DR PROSITE: P500237; G-PROTEIN_RECEPTOR; 1.
 CC PFM: PF00001; 7tm_1; 1.
 CC HSSP: P07700; IDEP.
 CC -----
 CC G-PROTEIN COUPLED RECEPTOR: TRANSMEMBRANE: GLYCOPROTEIN:
 CC MULTIGENE FAMILY: PHOSPHORYLATION: LIPOPROTEIN: PALMITATE.
 CC -----
 CC FT DOMAIN 1 25
 CC FT TRANSMEM 26 49
 CC FT DOMAIN 50 58
 CC FT TRANSMEM 59 77
 CC FT DOMAIN 78 97
 CC FT TRANSMEM 98 119
 CC FT DOMAIN 120 141
 CC FT TRANSMEM 142 164
 CC FT DOMAIN 165 189
 CC FT TRANSMEM 190 211
 CC FT DOMAIN 212 261
 CC FT TRANSMEM 262 283
 CC FT DOMAIN 284 294
 CC FT TRANSMEM 295 315
 CC FT DOMAIN 316 428
 CC FT CARBOHYD 8 8
 CC FT CARBOHYD 13 13
 CC FT DISULFID 96 175
 CC FT LIPID 329 329
 CC SO SEQUENCE 428 AA; 47398 MW; 8B794F0C CRC32;
 CC -----
 CC Query Match 65.4%; Score 51; DB 1; Length 428;
 CC Best Local Similarity 40.0%; Pred. No. 4.01e+00;
 CC Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC 122 YLAITAPLAY 131
 CC ||:||||:|
 CC QY 3 YLSTSSSLHY 12
 CC -----
 CC RESULT 5
 CC ID YJF8_YEAST STANDARD: PRT; 543 AA.
 CC AC P47041;
 CC DT 01-FEB-1996 (REL. 33, CREATED)
 CC DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 CC DT 01-NOV-1997 (REL. 33, LAST ANNOTATION UPDATE)
 CC DE HYPOTHETICAL 60.8 KD PROTEIN IN BTM1-PEP8 INTERGENIC REGION.
 CC GN J01058C OR J1141.
 CC OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).
 CC OC SACCAROMYCETES: ASCOMYCOTA: HEMIASCOCYCETES: SACCAROMYCETALES:
 CC OC SACCAROMYCETACEAE: SACCAROMYCETES.
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RA POHL T.M., ALJINOVIC G.;
 CC RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC CC -1- SIMILARITY: STRONG, TO YEAST YBR270C.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: 249333; G1008201; -
 CC DR HYPOTHETICAL PROTEIN.
 CC SO SEQUENCE 543 AA; 60840 MW; 8233FB93 CRC32;
 CC -----
 CC Query Match 65.4%; Score 51; DB 1; Length 543;
 CC Best Local Similarity 63.6%; Pred. No. 4.01e+00;
 CC Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC Db 182 ASPLDSSSLHY 192
 CC ||:||||:|
 CC QY 1 ASYLSTSSSLHY 11
 CC -----
 CC RESULT 6
 CC ID YGJR_ECOLI STANDARD: PRT; 328 AA.
 CC AC P42599; P42600; P76661;
 CC DT 01-NOV-1995 (REL. 32, CREATED)
 CC DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
 CC DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 CC DE HYPOTHETICAL 36.2 KD PROTEIN IN EBGC-UXXA INTERGENIC REGION.
 CC GN YGJR.
 CC OS ESCHERICHIA COLI.
 CC OC BACTERIA: PROTEOBACTERIA: GAMMA SUBDIVISION: ENTEROBACTERIACEAE;
 CC OC ESCHERICHIA.
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RP STRAIN-K12 / MG1655;
 CC RX MEDLINE: 97426617
 CC RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
 CC RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
 CC RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GORDEN M.A., ROSE D.J.,
 CC RA MAU B., SHAO Y.;
 CC RT "The complete genome sequence of Escherichia coli K-12."
 CC SCIENCE 277:1453-1474(1997).
 CC -1- SIMILARITY: BELONGS TO THE GPO/IDH/MOCA FAMILY. STRONG, TO
 CC B. SUBTILIS YULF.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U18997; G606025; ALT_FRAME.
 CC DR EMBL: U18997; G606026; ALT_FRAME.
 CC DR EMBL: AEO00390; G2367190; ALT_INIT.
 CC DR ECOGENE: EG12729; YGJR.
 CC KW HYPOTHETICAL PROTEIN: TRANSMEMBRANE.
 CC FT TRANSMEM 177 193
 CC SO SEQUENCE 328 AA; 36214 MW; 236E023A CRC32;
 CC -----
 CC Query Match 64.1%; Score 50; DB 1; Length 328;
 CC Best Local Similarity 41.7%; Pred. No. 6.49e+00;
 CC Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC Db 66 AVTIASPSLSLHY 77
 CC ||:||||:|
 CC QY 1 ASYLSTSSSLHY 12
 CC -----
 CC RESULT 7
 CC ID HPPD_ARATH STANDARD: PRT; 445 AA.
 CC AC P93836; O04330;
 CC DT 15-JUL-1998 (REL. 36, CREATED)
 CC DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
 CC DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE 4-HYDROXYPHENYLPYRUVATE DIOXYGENASE (EC 1.13.11.27) (4HPD) (HPD).
 GN HPD.
 OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
 OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
 CC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
 OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, MASSILEWSKJA;
 RA BARTLEY G.E., MAXWELL C.A., WITTENBACH V.A., SCOLNIK P.A.;
 RT "Cloning of an Arabidopsis thaliana cDNA for p-hydroxyphenylpyruvate
 dioxygenase."
 RL (IN) PLANT GENE REGISTER PGR97-065.
 RP [2]
 RC SEQUENCE FROM N.A.
 RA STRAIN-CV, COLUMBIA;
 RN NORRIS S.R., DELLAPENNA D.;
 RT SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RC SEQUENCE FROM N.A.
 RA STRAIN-CV, COLUMBIA;
 RN HSIEH T.-F., RODGERS M., MATRINGE M.;
 RT SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- CATALYTIC ACTIVITY: 4-HYDROXYPHENYLPYRUVATE + O(2) =
 HOMOCENTISATE + CO(2).
 CC -1- COFACTOR: IRON (BY SIMILARITY).
 CC -1- PATHWAY: PARTICIPATES IN THE CATABOLISM OF TYROSINE. IN THE
 CC -1- PATHWAY FOR BIOSYNTHESIS OF PRENYLOQUINONES.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE 4HPD FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U89267; G3392518; -
 DR EMBL: AF000228; G2145039; -
 DR EMBL: AF047834; G3098559; -
 KW OXIDOREDUCTASE; DIOXYGENASE; IRON.
 SQ SEQUENCE 445 AA; 48816 MW; B88784CA CRC32;
 Query Match 64.1%; Score 50; DB 1; Length 445;
 Best Local Similarity 58.3%; Pred. No. 6.49e+00;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 OY 86 ASYLSTSGDLRF 97
 1 ASYLSTSSSLHY 12

RL MOL. GEN. GENET. 244:151-159(1994).
 RN [2]
 RP PROTEIN SPLICING.
 RX MEDLINE: 97277324.
 RA WANG S., LIU X.-Q.;
 RT "Identification of an unusual intein in chloroplast ClpP protease of
 RT Chlamydomonas eugametos."
 RL J. BIOL. CHEM. 272:11869-11873(1997).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS TO SMALL PEPTIDES IN
 CC THE PRESENCE OF ATP AND MAGNESIUM. ALPHA-CASEIN IS THE USUAL TEST
 CC SUBSTRATE. IN THE ABSENCE OF ATP, ONLY OLIGOPEPTIDES SHORTER THAN
 CC FIVE RESIDUES ARE CLEAVED (SUCH AS SUCCINYL-LEU-TYR-1-NHMC, AND
 CC LEU-TYR-LEU-1-TYR-TRP, IN WHICH THE CLEAVAGE OF THE -TYR-1-LEU-
 CC AND -TYR-1-TRP- BOND ALSO OCCURS).
 CC -1- DOMAIN: THIS GENE CONTAINS TWO LARGE INSERTION SEQUENCES (IS1 AND
 CC CEU CLP INTEN1) THAT DIVIDE THE CLP GENE INTO THREE SEQUENCE
 CC DOMAINS. EACH INSERTION SEQUENCE FORMS A CONTINUOUS OPEN READING
 CC FRAME WITH ITS UPSTREAM AND DOWNSTREAM SEQUENCE DOMAINS.
 CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
 CC A POST-TRANSLATIONAL EXCISION OF THE THE INTERVENING REGION
 CC (INTEN1) FOLLOWED BY PEPTIDE LIGATION.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S14; ALSO KNOWN AS CLP
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L29402; G575471; -
 DR PROSITE: PS00381; CLP_PROTEASE_SER; 1.
 DR PROSITE: PS00382; CLP_PROTEASE_HIS; 1.
 DR PROSITE: PS00881; PROTEIN_SPLICING; FALSE_NEG.
 DR PFAM: PF00574; CLP_Protease; 3.
 DR MENDEL: 2294; CHLPP; 1.
 KW HYDROLASE; SERINE PROTEASE; CHLOROPLAST; PROTEIN SPLICING.
 FT CHAIN 1 447
 FT CHAIN 448 903
 FT CHAIN 904 1010
 FT DOMAIN 60 377
 FT ACT_SITE 419 419
 FT ACT_SITE 444 444
 FT ACT_SITE 444 444
 SQ SEQUENCE 1010 AA; 114551 MW; D08E3DA6 CRC32;
 Query Match 64.1%; Score 50; DB 1; Length 1010;
 Best Local Similarity 40.0%; Pred. No. 6.49e+00;
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 OY 951 YLTATETIHY 960
 3 YLTSSSLHY 12

RESULT 9
 ID PLTP_HUMAN STANDARD; PRT; 493 AA.
 AC P55058;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE PHOSPHOLIPID TRANSFER PROTEIN PRECURSOR (LIPID TRANSFER PROTEIN II).
 GN PLTP.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-27 AND 163-184.
 RC TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;
 RX MEDLINE: 94179366.
 RA DAY J.R., ALBERS J.J., LOFTON-DAY C.E., GILBERT T.L., CHING A.F.T.,
 RA GRANT F.J., O'HARA P.J., MARCOVINA S.M., ADOLPHSON J.L.;

Best Local Similarity 58.3%; Pred. No. 1.04e+01;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 883 ASYSLSTSLHY 894
QY 1 ASYSLSTSLHY 12

RESULT 12
ID DHSJ_BOVIN STANDARD; PRT; 158 AA.

AC 095123;
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE SUCCINATE DEHYDROGENASE (UBIQUINONE) CYTOCHROME B SMALL SUBUNIT
DE PRECURSOR (CPS3) (SUCCINATE-UBIQUINONE REDUCTASE MEMBRANE ANCHOR
DE SUBUNIT) (CPS3).
GN SDHD OR SDH4.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ACTIOACTYLIA; RUMINANTIA; PECORA; BOVIDEA; BOVINAE; BOS.
[1]
RP SEQUENCE FROM N.A.

RA SHENOI S.K., YU L., YU C.A.;
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SUBUNIT: COMPOSED OF A 27 KD IRON PROTEIN (IP), A 70 KD
CC FLAVOPROTEIN (FP) AND A CYTOCHROME B COMPOSED OF TWO INTEGRAL
CC MEMBRANE PROTEINS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U50987; G1575011.
KM TRICARBOXYLIC ACID CYCLE: ELECTRON TRANSPORT; HEME; TRANSMEMBRANE;
KW MITOCHONDRION; TRANSIT PEPTIDE.
FT TRANSIT 1 55
FT CHAIN 56 158
FT TRANSMEM 70 90
FT TRANSMEM 125 141
FT SEQUENCE 158 AA; 17096 MW; 703D5238 CRC32;

Query Match 61.5%; Score 48; DB 1; Length 158;
Best Local Similarity 41.7%; Pred. No. 1.67e+01;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 81 AAYLPCSAADY 92
QY 1 ASYSLSTSLHY 12

RESULT 13
ID DHSJ_HUMAN STANDARD; PRT; 159 AA.

AC 014521;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE SUCCINATE DEHYDROGENASE (UBIQUINONE) CYTOCHROME B SMALL SUBUNIT
DE PRECURSOR (CPS3) (SUCCINATE-UBIQUINONE REDUCTASE MEMBRANE ANCHOR
DE SUBUNIT).
GN SDHD OR SDH4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
[1]

RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;
RX MEDLINE: 98194224.
RA HIRAWAKE H., TANIWAKI M., KIJIMA S., KITA K.;
RT "Cytochrome b in human complex II (succinate-ubiquinone
RT oxidoreductase): cDNA cloning of the components in liver mitochondria
RT and chromosome assignment of the genes for the large (SDHC) and small
RT (SDHD) subunits to 1q21 and 11q23."
RL CYTOGENET. CELL GENET. 79:132-138(1997).
CC -1- SUBUNIT: COMPOSED OF A 27 KD IRON PROTEIN (IP), A 70 KD
CC FLAVOPROTEIN (FP) AND A CYTOCHROME B COMPOSED OF TWO INTEGRAL
CC MEMBRANE PROTEINS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AB006202; D1022913.
DR MIM: 602690.
KM TRICARBOXYLIC ACID CYCLE: ELECTRON TRANSPORT; HEME; TRANSMEMBRANE;
KW MITOCHONDRION; TRANSIT PEPTIDE.
FT TRANSIT 1 56
FT CHAIN 57 159
FT TRANSMEM 71 91
FT TRANSMEM 126 142
FT SEQUENCE 159 AA; 17043 MW; F4221825 CRC32;

Query Match 61.5%; Score 48; DB 1; Length 159;
Best Local Similarity 41.7%; Pred. No. 1.67e+01;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 82 AAYLPCSAADY 93
QY 1 ASYSLSTSLHY 12

RESULT 14
ID B3AR_CAVPO STANDARD; PRT; 351 AA.
AC 060483;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE BETA-3 ADRENERGIC RECEPTOR (FRAGMENT).
GN ADRB3.

OS CAVIA PORCELLUS (GUINEA PIG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; HYSTRICOGNATHI; CAVIIDAE; CAVIA.
[1]

RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE: 97151378.
RA ARGIE C., TAVERNIER G., D'ALLAIRE F., BENGTSSON T., MARTI L.,
RA CARPENE C., LAFONTAN M., BUKOWIECKI L.J., LANGIN D.;
RT "beta 3-adrenoceptor in guinea pig brown and white adipocytes: low
RT expression and lack of function."
RL AM. J. PHYSIOL. 271:R1729-R1738(1996).
CC -1- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE ACTION OF G
CC INDUCED ACTIVATION OF ADENYLYL CYCLASE THROUGH THE ACTION OF G
CC PROTEINS. BETA-3 IS INVOLVED IN THE REGULATION OF LIPOLYSIS AND
CC THERMOGENESIS.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: WHITE AND BROWN ADIPOSE TISSUES.
CC -1- THE GUINEA PIG DIFFERS FROM OTHER RODENTS BY AN ABSENCE OF BETA-3
CC ADRENERGIC EFFECTS AND BY LOW EXPRESSION IN BROWN AND WHITE
CC ADIPOSE TISSUES. IT IS CLOSER TO HUMAN OR PRIMATE THAN RODENT
CC BETA-3.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U51098; G1256416; -
 CC GCRDB: GCR_1166; -
 CC PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 CC PFAM: PF00001; 7tm_1; 1.
 CC G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 CC MULTIGENE FAMILY.
 CC
 CC FT DOMAIN 1 36
 CC FT TRANSMEM 37 60
 CC FT DOMAIN 61 69
 CC FT TRANSMEM 70 88
 CC FT DOMAIN 89 108
 CC FT TRANSMEM 109 130
 CC FT DOMAIN 131 152
 CC FT TRANSMEM 153 175
 CC FT DOMAIN 176 200
 CC FT TRANSMEM 201 222
 CC FT DOMAIN 223 290
 CC FT TRANSMEM 291 312
 CC FT DOMAIN 313 324
 CC FT TRANSMEM 325 345
 CC FT DOMAIN 346 >351
 CC FT DISULFID 107 186
 CC FT CARBOHYD 8 26
 CC FT NON_TER 351 351
 CC FT SEQUENCE 351 AA; 37364 MW; 72E17433 CRC32;
 CC
 CC Query Match 61.5%; Score 48; DB 1; Length 351;
 CC Best Local Similarity 40.0%; Pred. No. 1.67e+01;
 CC Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC Db 133 YLAVTNPLRY 142
 CC ||:||||:
 CC QY 3 YLSTSSSLHY 12
 CC
 CC RESULT 15
 CC B3AR_MOUSE STANDARD; PRT; 400 AA.
 CC P25962;
 CC 01-MAY-1992 (REL. 22, CREATED)
 CC 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 CC 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 CC DE BETA-3 ADRENERGIC RECEPTOR.
 CC GN ADRB3 OR ADRB3R.
 CC MUS MUSCULUS (MOUSE).
 CC OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 CC OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-SWISS;
 CC RX MEDLINE: 92037534.
 CC RA NAHMATAS C., BLIN N., ELALOUF J.M., MATTEI M.-G., STROSBERG A.D.,
 CC RA EMORINE L.J.;
 CC RT "Molecular characterization of the mouse beta 3-adrenergic receptor:
 CC RT relationship with the atypical receptor of adipocytes";
 CC RT EMBO J. 10:3721-3727(1991).
 CC RL [2]
 CC RN REVISIONS, SEQUENCE FROM N.A.
 CC RX MEDLINE: 93279311.
 CC RA VAN SPRONSEN A., NAHMATAS C., KRIEF S., BRIEND-SUTREN M.-M.,
 CC RA STROSBERG A.D., EMORINE L.J.;
 CC RT "The promoter and intron/exon structure of the human and mouse beta
 CC RT 3-adrenergic-receptor genes";

RL EUR. J. BIOCHEM. 213:1117-1124(1993).
 RN [3]
 RP SEQUENCE OF 378-400 FROM N.A.
 RC TISSUE-ADIPOSE TISSUE.
 RX MEDLINE: 93125503.
 RA GRANNEMAN J.G., LAHNERS K.N., RAO D.D.;
 RT "Rodent and human beta 3-adrenergic receptor genes contain an intron
 RT within the protein-coding block";
 RL MOL. PHARMACOL. 42:964-970(1992).
 CC -1- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
 CC INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
 CC PROTEINS. BETA-3 IS INVOLVED IN THE REGULATION OF LIPOLYSIS AND
 CC THERMOGENESIS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: WHITE AND BROWN ADIPOSE TISSUES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X72862; G298113; -
 CC EMBL: X60438; G50110; -
 CC EMBL: S53280; G263087; -
 CC PIR: S18548; S18548.
 CC PIR: S32804; S32804.
 CC GCRDB: GCR_0253; -
 CC GCRDB: GCR_0551; -
 CC GCRDB: GCR_0708; -
 CC MGI: 87939; ADRB3.
 CC PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 CC PFAM: PF00001; 7tm_1; 1.
 CC HSSP: P07700; IDEP.
 CC G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 CC MULTIGENE FAMILY; PHOSPHORYLATION; LIPOPROTEIN; PALMITATE.
 CC
 CC FT DOMAIN 1 36
 CC FT TRANSMEM 37 60
 CC FT DOMAIN 61 69
 CC FT TRANSMEM 70 88
 CC FT DOMAIN 89 108
 CC FT TRANSMEM 109 130
 CC FT DOMAIN 131 152
 CC FT TRANSMEM 153 175
 CC FT DOMAIN 176 200
 CC FT TRANSMEM 201 222
 CC FT DOMAIN 223 289
 CC FT TRANSMEM 290 311
 CC FT DOMAIN 312 323
 CC FT TRANSMEM 324 344
 CC FT DOMAIN 345 400
 CC FT CARBOHYD 8 26
 CC FT DISULFID 107 186
 CC FT LIPID 358 358
 CC FT SEQUENCE 400 AA; 43006 MW; 474A96AC CRC32;
 CC
 CC Query Match 61.5%; Score 48; DB 1; Length 400;
 CC Best Local Similarity 40.0%; Pred. No. 1.67e+01;
 CC Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC Db 133 YLAVTNPLRY 142
 CC ||:||||:
 CC QY 3 YLSTSSSLHY 12
 CC
 CC Search completed: Thu Sep 2 12:43:17 1999
 CC Job time : 7 secs.

THIS PAGE BLANK (USPTO)

 W P O S E R I E S
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
 on: Thu Sep 2 12:42:33 1999; Maspar time 3.14 Seconds
 152.929 Million cell updates/sec
 Modular output not generated.

Title: >US-08-599-226-32
 Description: (1-12) from US08599226.pep
 Perfect Score: 78
 Sequence: 1 ASYLSTSSSLHY 12

Scoring table:
 PAM 150
 Gap 15

Searched: 122810 segs, 40068593 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: PIR60
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 24.334; Variance 30.685; scale 0.793

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	59	75.6	989	2	I56333	apolipoprotein B - ra	2.24e+01
2	52	66.7	296	2	S45336	finger protein, Spl/e	6.21e+00
3	52	66.7	828	2	S34695	hypothetical protein	6.21e+00
4	51	65.4	303	2	S45461	hypothetical protein	9.75e+00
5	51	65.4	428	2	A55044	beta-4C-adrenergic re	9.75e+00
6	51	65.4	543	2	S56830	probable purine nucle	9.75e+00
7	51	65.4	2150	2	S71629	sensory transduction	9.75e+00
8	50	64.1	236	2	S46280	clip protein - Chlamy	1.52e+01
9	50	64.1	253	2	A71648	pseudouridylylate synth	1.52e+01
10	49	62.8	334	2	D65097	YgjR protein - Escher	1.52e+01
11	49	62.8	135	2	S24320	Ig kappa chain precur	2.36e+01
12	49	62.8	493	2	A53533	phospholipid transfer	2.36e+01
13	49	62.8	503	2	P64713	protein-export membra	2.36e+01
14	49	62.8	526	2	D71805	hypothetical protein	2.36e+01
15	49	62.8	968	2	T00353	hypothetical protein	2.36e+01
16	49	62.8	1392	2	T01908	hypothetical protein	2.36e+01
17	49	62.8	2410	1	J01948	genome polypeptide 1	2.36e+01
18	49	62.8	2412	1	J01537	genome polypeptide 1	2.36e+01
19	48	61.5	400	2	S32804	beta-3-adrenergic rec	3.63e+01
20	48	61.5	400	2	A41679	beta-3-adrenergic rec	3.63e+01
21	48	61.5	400	2	A53281	beta-3-adrenergic rec	3.63e+01
22	48	61.5	405	2	S65459	beta-3-adrenergic rec	3.63e+01
23	48	61.5	408	1	QRHUBE	beta-3-adrenergic rec	3.63e+01

24	48	61.5	414	1	QRHUB3	beta-3-adrenergic rec	3.63e+01
25	48	61.5	418	2	G02953	beta-3-adrenergic rec <td>3.63e+01</td>	3.63e+01
26	48	61.5	441	2	G70822	probable secy protein <td>3.63e+01</td>	3.63e+01
27	48	61.5	590	2	S72961	GTP synthase (EC 6.3.3) <td>3.63e+01</td>	3.63e+01
28	48	61.5	826	2	S36203	iron-responsive element <td>3.63e+01</td>	3.63e+01
29	48	61.5	952	2	B57238	iron-responsive element <td>3.63e+01</td>	3.63e+01
30	48	61.5	963	2	A57238	iron-responsive element <td>3.63e+01</td>	3.63e+01
31	48	61.5	1255	1	A24571	protein-tyrosine kinase <td>3.63e+01</td>	3.63e+01
32	48	61.5	1469	2	A55095	chromosome condensati <td>5.56e+01</td>	5.56e+01
33	47	60.3	105	2	S69755	hypothetical protein <td>5.56e+01</td>	5.56e+01
34	47	60.3	112	2	H64215	hypothetical protein <td>5.56e+01</td>	5.56e+01
35	47	60.3	119	2	S65534	light-harvesting chl <td>5.56e+01</td>	5.56e+01
36	47	60.3	265	2	S38380	Hrox1 protein - Calif <td>5.56e+01</td>	5.56e+01
37	47	60.3	291	1	S18718	transcription factor <td>5.56e+01</td>	5.56e+01
38	47	60.3	291	1	A31305	transcription factor <td>5.56e+01</td>	5.56e+01
39	47	60.3	325	2	S75747	hypothetical protein <td>5.56e+01</td>	5.56e+01
40	47	60.3	394	2	S23481	beta protein - Aclinet <td>5.56e+01</td>	5.56e+01
41	47	60.3	434	2	S42676	KEB1 protein - yeast <td>5.56e+01</td>	5.56e+01
42	47	60.3	469	2	D70048	ABC transporter (amin <td>5.56e+01</td>	5.56e+01
43	47	60.3	533	2	S62489	hypothetical protein <td>5.56e+01</td>	5.56e+01
44	47	60.3	684	2	G70744	hypothetical protein <td>5.56e+01</td>	5.56e+01
45	47	60.3	1746	2	S19694	tenascin precursor - <td>5.56e+01</td>	5.56e+01

ALIGNMENTS

RESULT 1
 ENTRY 1
 TITLE I56333 #type fragment
 ORGANISM apolipoprotein B - rat (fragment)
 DATE 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 23-Feb-1997

ACCESSIONS
 REFERENCE I56333
 #authors Reuben, M.A.; Svenson, K.L.; Doilittle, M.H.; Johnson, D.F.; Lusis, A.J.; Elverson, J.
 #journal J. Lipid Res. (1988) 29:1337-1347
 #title Biosynthetic relationships between three rat apolipoprotein B peptides.
 #cross-references MIMD:89176719
 #accession I56333
 #status preliminary; translated from GB/EMBL/DBJ
 #molecule_type mRNA
 #residues 1-989 #label RES
 #cross-references GB:M27440; NID:9623548; PID:9623549

GENETICS
 #gene apob
 #classification #superfamily apolipoprotein B
 #summary #length 989 #checksum 1918

Query Match 75.6%; Score 59; DB 2; Length 989;
 Best Local Similarity 70.0%; Pred. No. 2.24e+01;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 175 Y10ASTSLHY 184
 |||:|||||
 QY 3 YLSTSSSLHY 12

RESULT 2
 ENTRY 2
 TITLE S45336 #type complete
 ORGANISM finger protein, Spl/egr-like - fruit fly (Drosophila sp.)
 DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997

ACCESSIONS
 REFERENCE S45336
 #authors Broecker, G.; Chu-Lacraff, O.; Doe, C.Q.; Cohen, B.; Weigel, D.; Taubert, H.; Jaekle, H.
 #journal Nature (1994) 369:664-668
 #title Spl/egr-like zinc-finger protein required for endoderm specification and germ-layer formation in Drosophila.
 #accession S45336

##status preliminary
##molecule-type DNA
##residues 1-296 ##label BRO
SUMMARY #length 296 #molecular-weight 33598 #checksum 2521

Query Match 66.7%; Score 52; DB 2; Length 296;
Best Local Similarity 50.0%; Pred. No. 6.21e+00;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 127 APFLNASDLY 138
1 ASYLSTSSSLHY 12

RESULT 3
ENTRY S34695 #type complete
TITLE hypothetical protein YKL176c - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES hypothetical protein YKL642
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Nov-1997
ACCESSIONS S34695; S38008; S38006; S44583; S38403
REFERENCE S34679
#authors Wieman, S.; Voss, H.; Schwager, C.; Rupp, T.; Stegemann, J.; Zimmermann, J.; Grochues, D.; Sensen, C.; Erfle, H.; Hewitt, N.; Banrevi, A.; Ansoerge, W.
#submission Submitted to the EMBL Data Library, July 1993
#description Sequencing and analysis of 51.5 kilobases on the left arm of chromosome XI from Saccharomyces cerevisiae reveals 23 open reading frames including the FAS1 gene.

#accession S34695
##molecule-type DNA
##residues 1-828 ##label WIE
##cross-references EMBL:X74151; NID:g450365; PID:g395250
REFERENCE S37825
#authors Wiemann, S.; Voss, H.; Schwager, C.; Rupp, T.; Grochues, D.; Sensen, C.; Stegemann, J.; Zimmermann, J.; Erfle, H.; Hewitt, N.; Ansoerge, W.
#submission Submitted to the Protein Sequence Database, March 1994
#accession S38008
##molecule-type DNA
##residues 1-828 ##label W12
##cross-references EMBL:Z28176; NID:g486309; PID:g486310; MIPS:YKL176c
REFERENCE S37976
#authors Vandenhof, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
#submission Submitted to the Protein Sequence Database, March 1994
#accession S38006
##molecule-type DNA
##residues 1-306 ##label VAN
##cross-references EMBL:Z28176; MIPS:YKL176c
REFERENCE S44383
#authors Vandenhof, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
#journal Yeast (1994) 10:25-33
#title Sequencing and analysis of a 20.5 kb DNA segment located on the left arm of yeast chromosome XI.
#accession S44383
##status Translation not shown
##molecule-type DNA
##residues 1-306 ##label VA2
##cross-references EMBL:Z28678; NID:g407503; PID:g407504
##experimental_source strain S288C

GENETICS
#map_position 11L
SUMMARY #length 828 #molecular-weight 93358 #checksum 6781

Query Match 66.7%; Score 52; DB 2; Length 828;
Best Local Similarity 66.7%; Pred. No. 6.21e+00;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 483 SFSKSGSSSLHY 494
1 ASYLSTSSSLHY 12

RESULT 4
ENTRY S45461 #type complete
TITLE hypothetical protein YNL006w - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES hypothetical protein N2005
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 13-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 17-Mar-1999
ACCESSIONS S45461; S62915; S62917; S45123
REFERENCE S45456
#authors Verhasselt, P.; Aert, R.; Voet, M.; Volckaert, G.
#journal Yeast (1994) 10:945-951
#title Nucleotide sequence analysis of an 8887 bp region of the left arm of yeast chromosome XIV, encompassing the centromere sequence.

#cross-references MUID:95076713
#accession S45461
##status Translation not shown
##molecule-type DNA
##residues 1-303 ##label VER
##cross-references EMBL:X77114; NID:g496710; PID:g496716
REFERENCE S62910
#authors Aert, R.; Verhasselt, P.; Voet, M.; Volckaert, G.
#submission Submitted to the Protein Sequence Database, April 1996
#accession S62915
##molecule-type DNA
##residues 1-303 ##label AER
##cross-references EMBL:Z71282; NID:g1301821; PID:e239859; PID:g1301822; MIPS:YNL006w
REFERENCE S62916
#authors Dolignon, F.; Crouzet, M.
#submission Submitted to the Protein Sequence Database, April 1996
#accession S62917
##molecule-type DNA
##residues 1-303 ##label DOI
##cross-references EMBL:Z71282; NID:g1301821; PID:e239859; PID:g1301822; MIPS:YNL006w
##experimental_source strain S288C

GENETICS
#map_position 14L
CLASSIFICATION #superfamily WD repeat homology
FEATURE
71-104 #domain WD repeat homology #label WD2\
112-145 #domain WD repeat homology #label WD3\
203-236 #domain WD repeat homology #label WD5\
246-279 #domain WD repeat homology #label WD6
SUMMARY #length 303 #molecular-weight 34034 #checksum 537

Query Match 65.4%; Score 51; DB 2; Length 303;
Best Local Similarity 58.3%; Pred. No. 9.75e+00;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 262 SAYLVYASSDHY 273
1 ASYLSTSSSLHY 12

RESULT 5
ENTRY A55044 #type complete
TITLE beta-4C-adrenergic receptor - turkey
ORGANISM #formal_name Melalegris gallipavo #common_name common turkey
DATE 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 08-Sep-1997
ACCESSIONS A55044
REFERENCE A55044

#authors Chen, X.; Harden, T.K.; Nicholas, R.A.
#journal J. Biol. Chem. (1994) 269:24810-24819
#title Molecular cloning and characterization of a novel
beta-adrenergic receptor.
#accession A55044
#status preliminary
##molecule-type DNA
##residues 1-428 ##label CHE
##cross-references GB:U13978; NID:9555881; PID:9555882
GENETICS
#intons 416/2
CLASSIFICATION #superfamily vertebrate rhodopsin
KEYWORDS neurotransmitter receptor; transmembrane protein
SUMMARY #length 428 #molecular-weight 47398 #checksum 8085
Query Match 65.4%; Score 51; DB 2; Length 428;
Best Local Similarity 40.0%; Pred. No. 9.75e+00;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
122 YLAIAPLOY 131
||: ||: ||: ||
3 YLSTSSSLH 12
RESULT 6
ENTRY S56830 #type complete
TITLE probable purine nucleotide-binding protein YJL058c - yeast
ALTERNATE_NAMES (saccharomyces cerevisiae)
ORGANISM probable membrane protein YJL058c; protein J1141
#formal_name Saccharomyces cerevisiae
DATE 05-May-1995 #sequence_revision 08-Sep-1995 #text_change
14-Nov-1997
ACCESSIONS S56830
REFERENCE S56793
#authors Pohl, T.M.; Ajjino, G.
#submission submitted to the Protein Sequence Database, September 1995
#accession S56830
##molecule-type DNA
##residues 1-543 ##label NOV
##cross-references EMBL:249333; NID:91008200; PID:91008201; MIPS:YJL058c
GENETICS
#map_position 10L
KEYWORDS transmembrane protein
SUMMARY #length 543 #molecular-weight 60840 #checksum 5268
Query Match 65.4%; Score 51; DB 2; Length 543;
Best Local Similarity 63.6%; Pred. No. 9.75e+00;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
182 ASFLDSTLH 192
||: ||: ||: ||
1 ASYLTSSSLH 11
OY
RESULT 7
ENTRY S71629 #type complete
TITLE sensory transduction histidine kinase dhka - slime mold
ORGANISM (Dictyostellium discoideum)
#formal_name Dictyostellium discoideum
DATE 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change
24-Sep-1998
ACCESSIONS S71629
REFERENCE S71629
#authors Wang, N.; Shauly, G.; Escalante, R.; Loomis, W.F.
#journal EMBO J. (1996) 15:3890-3898
#title A two-component histidine kinase gene that functions in
Dictyostellium development.
#cross-references MUID:96324397
#accession S71629
#status nucleic acid sequence not shown
##molecule-type mRNA
##residues 1-2150 ##label WAN
##cross-references EMBL:U42597

##experimental_source strain Ax4
GENETICS
#gene dhka
#map_position 6
CLASSIFICATION #superfamily response regulator homology
KEYWORDS autophosphorylation; phosphoprotein; phosphotransferase;
two-component regulatory system
FEATURE
2027-2142
2076 #domain response regulator homology #label RRH
#binding_site phosphate (Asp) (covalent) #status
predicted
SUMMARY #length 2150 #molecular-weight 239635 #checksum 5691
Query Match 65.4%; Score 51; DB 2; Length 2150;
Best Local Similarity 70.0%; Pred. No. 9.75e+00;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 180 NYLNSSSLH 189
||: ||: ||: ||
2 YLSTSSSLH 11
OY
RESULT 8
ENTRY S46280 #type complete
TITLE clp protein - Chlamydomonas eugametos chloroplast
ORGANISM #formal_name Chloroplast Chlamydomonas eugametos
DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
01-Aug-1997
ACCESSIONS S46280
REFERENCE S46279
#authors Huang, C.; Wang, S.; Chen, L.; Lemieux, C.; Otis, C.; Turmel,
M.; Liu, X.O.
#journal Mol. Gen. Genet. (1994) 244:151-159
#title The Chlamydomonas chloroplast clp gene contains translated
large insertion sequences and is essential for cell growth.
#cross-references MUID:94329067
#accession S46280
##status preliminary
##molecule-type DNA
##residues 1-236 ##label HUA
GENETICS
#genome chloroplast
KEYWORDS chloroplast
SUMMARY #length 236 #molecular-weight 26638 #checksum 3940
Query Match 64.1%; Score 50; DB 2; Length 236;
Best Local Similarity 40.0%; Pred. No. 1.52e+01;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Db 177 YLATEITHY 186
||: ||: ||: ||
3 YLSTSSSLH 12
OY
RESULT 9
ENTRY A71648 #type complete
TITLE pseudouridylylate synthase I (trua) RP857 - Rickettsia
prowazekii
ORGANISM #formal_name Rickettsia prowazekii
DATE 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change
21-Nov-1998
ACCESSIONS A71648
REFERENCE A71630
#authors Andersson, S.G.E.; Zomrodipour, A.; Andersson, J.O.;
Sicheritz-Ponten, T.; Almark, U.C.M.; Podowski, R.M.;
Neeslund, A.K.; Eriksson, A.S.; Winkler, H.H.; Kurland,
C.G.
#journal Nature (1998) 396:133-140
#title The genome sequence of Rickettsia prowazekii and the origin
of mitochondria.
#accession A71648
#status preliminary; nucleic acid sequence not shown;
translation not shown

```

##molecule_type DNA
##residues 1-253 ##label AND
##cross-references GB:A235273; GB:A235269; NID:g3861237; PID:e1343126;
PID:g3861382
##experimental_source strain Madrid E

GENETICS
#gene trna: RP857
SUMMARY #length 253 #molecular_weight 29075 #checksum 2010

Query Match 64.1% Score 50; DB 2; Length 253;
Best Local Similarity 60.0% Pred. No. 1.52e+01;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 194 YISAPFLHY 203
|:|:|:|
QY 3 YLSTSSSLHY 12

RESULT 10
ENTRY #type complete
TITLE yjgR protein - Escherichia coli (strain K-12)
ORGANISM #formal_name Escherichia coli
DATE 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change
13-Sep-1998
ACCESSIONS D65097
REFERENCE D65097
AUTHORS Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.;
Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.;
Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.M.;
Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao,
Y.
Science (1997) 277:1453-1462
The complete genome sequence of Escherichia coli K-12.
#cross-references MUID:97426617
#accession D65097
#status preliminary: nucleic acid sequence not shown;
translational not shown

GENETICS
##experimental_source strain K-12, substrain MG1655

#gene yjgR
CLASSIFICATION #superfamily Escherichia coli yjgR protein
SUMMARY #length 334 #molecular_weight 36990 #checksum 2682

Query Match 64.1% Score 50; DB 2; Length 334;
Best Local Similarity 41.7% Pred. No. 1.52e+01;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 72 AVYIAPNSLHF 83
|:|:|:|
QY 1 ASYLSTSSSLHY 12

RESULT 11
ENTRY #type complete
TITLE Ig kappa chain precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 02-Dec-1993 #sequence_revision 17-Nov-1995 #text_change
08-Sep-1997
ACCESSIONS S24320
REFERENCE S24319
AUTHORS Aucuttier, P.; Khamilchi, A.A.; Preud'homme, J.L.; Bauwens,
M.; Touchard, G.; Cogne, M.;
Biochem J. (1992) 285:149-152
Complementary DNA sequence of human amyloidogenic
immunoglobulin light-chain precursors.
#cross-references MUID:92344562
#accession S24320
#status preliminary
#molecule_type mRNA

```

```

#residues      1-135 #label AUC
##cross-references EMBL:X61133; NID:g32810.; PID:g32811
##note          The authors translated the codon CAA for residue 122 as
                  Gln
CLASSIFICATION  #superfamily immunoglobulin V region; immunoglobulin homology
KEYWORDS        heterotetramer; immunoglobulin
SUMMARY         #length 135 #molecular-weight 148k4 #checksum 5727

Query Match     62.8%; Score 49; DB 2; Length 135;
Best Local Similarity 50.0%; Pred. No. 2,36e+01;
Matches         6; Conservative         4; Mismatches    2; Indels    0; Gaps   0;

Db              73 ASYLGTGVPLRF 84
               |||||:::||:
OY             1 ASYLSTSSSLHY 12

RESULT 12
ENTRY      A53533 #type complete
TITLE      phospholipid transfer protein precursor - human
ALTERNATE_NAMES
ORGANISM   Homo sapiens #common_name man
DATE       27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change
           17-Mar-1999
ACCESSIONS A53533
REFERENCE   A53533
AUTHORS     Day, J.R.; Albers, J.J.; Lofton-Day, C.E.; Gilbert, T.L.;
            Chling, A.F.T.; Grant, F.J.; O'Hara, P.O.; Marcovina, S.M.;
            Adolphson, J.L.
            J. Biol. Chem. (1994) 269:9388-9391
            Complete cDNA encoding human phospholipid transfer protein
            from human endothelial cells.
#cross-references MIMD:94179366
#accession   A53533
#status      preliminary
#molecule_type mRNA
##residues   1-493 ##label DAY
##cross-references GB:L26232; NID:g468325; PID:g468326
GENETICS
#gene        GDB:340911; OMIM:172425
##cross-references GDB:20pter-20qter
KEYWORDS     map_position 20pter-20qter
FEATURE      glycoprotein; phosphoprotein
1-17         #domain signal sequence #status predicted #label SIG\
18-493       #product phospholipid transfer protein #status
                experimental #label MAT
SUMMARY      #length 493 #molecular-weight 54739 #checksum 6063

Query Match     62.8%; Score 49; DB 2; Length 493;
Best Local Similarity 33.3%; Pred. No. 2,36e+01;
Matches         4; Conservative         6; Mismatches    2; Indels    0; Gaps   0;

Db              453 AGFLTGADLHF 464
               |::|:::||:
OY             1 ASYLSTSSSLHY 12

RESULT 13
ENTRY      F64713 #type complete
TITLE      protein-export membrane protein - Helicobacter pylori (strain
           26695)
ORGANISM     #formal_name Helicobacter pylori
DATE         09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change
           12-Feb-1999
ACCESSIONS     F64713
REFERENCE      A64520
AUTHORS        Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.;
            Sutton, G.G.; Fleischmann, R.D.; Ketchum, K.A.; Klenk,
            H.P.; Gill, S.; Dougherty, B.A.; Nelson, K.; Quackenbush,
            J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus, B.;
            Richardson, D.; Dodson, R.; Khairik, H.G.; Glodek, A.;
            McKenney, K.; Fitzgerald, L.M.; Lee, N.; Adams, M.D.;

```

Hikey, E.K.; Berg, D.E.; Gocayne, J.D.; Utterback, T.R.;
Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Meliman, J.M.;
Fuji, C.; Bowman, C.; Matthey, L.; Wallin, E.; Hayes,
W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
C.M.; Venter, J.C.
Nature (1997) 388:539-547
#journal The complete genome sequence of the gastric pathogen
#title Helicobacter pylori.
#cross-references MUID:97394467
#accession F64713
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-503 #label TOM
#cross-references GB:AE000552; GB:AE000511; NID:g2314720; PID:g2314730;
TIGR:HP1550

GENETICS
#start_codon GTG
CLASSIFICATION #superfamily protein export membrane protein secd
#length 503 #molecular_weight 54247 #checksum 3320
SUMMARY

Query Match 62.8%; Score 49; DB 2; Length 503;
Best Local Similarity 60.0%; Pred. No. 2.36e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 38 YL5LASALEY 47
||| :|||
QY 3 YL5TSSSLHY 12

RESULT 14
ENTRY D71805 #type complete
TITLE protein-export membrane protein - Helicobacter pylori (strain
J99)
ORGANISM #formal_name Helicobacter pylori
#strain J99
DATE 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
05-Mar-1999
ACCESSIONS D71805
REFERENCE A71800
#authors Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.;
Doig, P.C.; Smith, D.R.; Noonan, B.; Guild, B.C.; deJonge,
B.L.; Carmel, G.; Tummino, P.J.; Caruso, A.;
Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.;
Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Voyis,
G.F.; Trust, T.J.
Nature (1999) 397:176-180
#journal Genomic sequence comparison of two unrelated isolates of the
#title human gastric pathogen Helicobacter pylori.
#cross-references MUID:99120557
#accession D71805
#status preliminary
#molecule_type DNA
#residues 1-526 #label ARN
#cross-references GB:AE001567; GB:AE001439; NID:g4156065; PID:g4156069
#experimental_source strain J99

GENETICS
#gene secd
CLASSIFICATION #superfamily protein export membrane protein secd
SUMMARY #length 526 #molecular_weight 56796 #checksum 5813

Query Match 62.8%; Score 49; DB 2; Length 526;
Best Local Similarity 60.0%; Pred. No. 2.36e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 60 YL5LASALEY 69
||| :|||
QY 3 YL5TSSSLHY 12

RESULT 15
ENTRY T00353 #type fragment
TITLE hypothetical protein KIAA0697 - human (fragment)

ORGANISM #formal_name Homo sapiens #common_name man
DATE 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change
01-Feb-1999
ACCESSIONS T00353
REFERENCE Z14142
#authors Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka,
A.; Kotani, H.; Nomura, N.; Ohara, O.
DNA Res. (1998) 5:169-176
#journal Prediction of the coding sequences of unidentified human
#title genes. X. The complete sequences of 100 new cDNA clones
from brain which can code for large proteins in vitro.
#accession T00353
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-968 #label ISH
#cross-references EMBL:AB014597; NID:d1204352; PID:d1032633
#experimental_source brain; clone HK04486

GENETICS
#note KIAA0697
SUMMARY #length 968 #checksum 7201

Query Match 62.8%; Score 49; DB 2; Length 968;
Best Local Similarity 54.5%; Pred. No. 2.36e+01;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 653 SSYLPNSDPLH 663
||| :|||
QY 1 ASYLTSSSLH 11

Search completed: Thu Sep 2 12:42:53 1999
Job time : 20 secs.

THIS PAGE BLANK (USPTO)

 WISEMAN (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSPRCH_PP protein - protein database search, using Smith-Waterman algorithm

On: Thu Sep 2 12:43:36 1999; Maspar time 4.40 Seconds 148.939 Million cell updates/sec

Similar output not generated.

Title: >US-08-599-226-32
 Description: (1-12) from US08599226.pep
 Perfect Score: 78
 Sequence: 1 ASYLSTSSSLHY 12

Scoring table:
 PAM 150
 Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
 5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_orcanelle
 9:sp_phage 10:sp_plant 11:sp_protent 12:sp_unclassified
 13:sp_vertebrate 14:sp_virus

Statistics: Mean 24.038; Variance 31.822; scale 0.755

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	59	75.6	989	11	063052	APOLIPOPROTEIN B (FRAG	4.42e+01	
2	52	66.7	296	5	Q26364	HKB-SEGMENTATION GAP-G	1.08e+01	
3	52	66.7	372	5	P91143	SIMILAR TO ACETYLTRANS	1.08e+01	
4	52	66.7	531	8	O47582	NADH DEHYDROGENASE SUB	1.08e+01	
5	52	66.7	1338	5	O61198	FI5E6.6 PROTEIN.	1.08e+01	
6	51	65.4	236	6	O77542	OLFACTORY RECEPTOR (FR	1.68e+01	
7	51	65.4	238	6	O77543	OLFACTORY RECEPTOR (FR	1.68e+01	
8	51	65.4	321	11	O70265	OLFACTORY RECEPTOR-LIK	1.68e+01	
9	51	65.4	321	11	O70267	OLFACTORY RECEPTOR-LIK	1.68e+01	
10	51	65.4	1494	11	O88902	PROTEIN TYROSINE PHOSP	1.68e+01	
11	51	65.4	2150	5	O23863	HISTIDINE KINASE A.	1.68e+01	
12	51	65.4	2219	5	O23388	ZK1067.2 PROTEIN.	1.68e+01	
13	50	64.1	220	13	P70013	OLFACTORY RECEPTOR (FR	2.58e+01	
14	50	64.1	357	5	O45973	YE2A.6 PROTEIN.	2.58e+01	
15	50	64.1	419	10	O82449	P-HYDROXYPHENYL-PYRUVAT	2.58e+01	
16	50	64.1	444	5	P91141	SIMILAR TO ACETYLTRANS	2.58e+01	
17	50	64.1	649	3	O60167	PROTEIN COMPLEX ASSEMB	2.58e+01	
18	49	62.8	246	5	Q22895	SIMILAR TO FAMILY 1 OF	3.93e+01	
19	49	62.8	503	2	O26074	PROTEIN-EXPORT MEMBRAN	3.93e+01	
20	49	62.8	809	5	Q20702	SIMILAR TO MATRIN F/G.	3.93e+01	

21	49	62.8	968	4	O75179	KIAA0697 PROTEIN (FRAG	3.93e+01
22	49	62.8	1392	10	O82493	T12H20.12 PROTEIN.	3.93e+01
23	48	61.5	176	9	O37837	EXPRESSED IN UNINDUCED	5.97e+01
24	48	61.5	176	8	O36294	CYTOCHROME B (FRAGMENT	5.97e+01
25	48	61.5	204	3	O12097	HYPOTHETICAL 22.6 KD P	5.97e+01
26	48	61.5	301	3	O22299	T07C5.4 PROTEIN.	5.97e+01
27	48	61.5	349	5	O17959	M01B2.5 PROTEIN.	5.97e+01
28	48	61.5	910	3	O59796	PUTATIVE RECEPTOR ASSO	5.97e+01
29	48	61.5	1046	2	P96156	CHITODEXTRINASE.	5.97e+01
30	48	61.5	1406	4	O15082	KIAA0377.	5.97e+01
31	47	60.3	98	6	O18812	DELTA 3 PIT-1 (FRAGMEN	5.00e+01
32	47	60.3	122	10	P93302	ORF122A.	9.00e+01
33	47	60.3	185	10	O38688	LIGHT-HARVESTING CHLOR	9.00e+01
34	47	60.3	198	4	O75805	HXA-9A.	9.00e+01
35	47	60.3	265	5	O25144	HXA-9A.	9.00e+01
36	47	60.3	285	6	O02706	PIT-1, PARTIAL (FRAGME	9.00e+01
37	47	60.3	435	13	P70033	NUCLEAR ORPHAN RECEPTO	9.00e+01
38	47	60.3	463	5	O22925	COSMID C50B3.	9.00e+01
39	47	60.3	613	3	O74954	HYPOTHETICAL 67.2 KD P	9.00e+01
40	47	60.3	646	14	O65281	MAJOR CAPSID PROTEIN.	9.00e+01
41	47	60.3	725	2	O52978	PHAI1.B.C.D.E.F.G GEN	9.00e+01
42	47	60.3	932	5	O01623	SIMILAR TO LIGAND-GATE	9.00e+01
43	47	60.3	1101	5	O22378	CODED FOR BY C. ELEGAN	9.00e+01
44	47	60.3	1746	6	O29116	TENASCIN PRECURSOR (TN	9.00e+01
45	47	60.3	1943	5	O62341	R0F6.8B PROTEIN.	9.00e+01

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	989 AA.
AC	063052:			
DT	01-NOV-1996 (TREMBLREL. 01, CREATED)			
DT	01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)			
DT	01-NOV-1996 (TREMBLREL. 08, LAST ANNOTATION UPDATE)			
DE	APOLIPOPROTEIN B (FRAGMENT).			
GN	APOB.			
OS	RATTUS NORVEGICUS (RAT).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;			
OC	SCIROGNATHI; MORINAE; MORINAE; RATTUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-LIVER:			
RX	MEDLINE: 89176719.			
RA	REUSEN M.A., SVENSON K.L., DOOLITTLE M.H., JOHNSON D.F., LUSIS A.J.,			
RA	ELOVSON J.:			
RT	"Biosynthetic relationships between three rat apolipoprotein B			
RT	peptides."			
RL	J. LIPID RES. 29:1337-1347(1988).			
DR	EMBL: M27440; G623549; -.			
KW	LIPOPROTEIN.			
FT	NON_TER			
SQ	SEQUENCE 989 AA; 113305 MW; 4EFC0E6E CRC32;			
Query Match	75.6%;			
Best Local Similarity	70.0%;			
Matches	7; Conservative			
	2; Mismatches			
	1; Indels			
	0; Gaps			
	0;			
DB	175 YLOASTSLHY 184			
OY	3 YLSTSSSLHY 12			
RESULT	2			
ID	Q26364:	PRELIMINARY:	PRT:	296 AA.
AC	Q26364:			
DT	01-NOV-1996 (TREMBLREL. 01, CREATED)			
DT	01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)			
DT	01-NOV-1996 (TREMBLREL. 08, LAST ANNOTATION UPDATE)			
DE	HKB-SEGMENTATION GAP-GENE HUCKERBEIN.			
GN	HKB.			
OS	DROSOPHILA MELANOGASTER (FRUIT FLY).			
OC	EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;			

OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
 OC DROSOPHILIDE; DROSOPHILA.
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 94268555.
 RA BRONNER G., CHU-LAGRAFF Q., DOE C.O., COHEN B., WEIGEL D., TAUBERT H.,
 RA JACKIE H.,
 RT "Sp/egf-like zinc-finger protein required for endoderm specification
 and germ-layer formation in Drosophila.";
 RL NATURE 369:664-668(1994).
 DR EMBL: S71230; G547124; -
 DR FLYBASE: FBgn0001204; hkd.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 3.
 DR PFAM: PF00096; zf-C2H2; 3.
 KW ZINC-FINGER; METAL-BINDING; DNA-BINDING.
 SQ SEQUENCE 296 AA; 33598 MW; 5716DF5 CRC32;

Query Match 66.7%; Score 52; DB 5; Length 296;
 Best Local Similarity 50.0%; Pred. No. 1.08e+01;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 DB 127 APFLSASDLY 138
 QY 1 ASYSTSSSLHY 12

RESULT 3
 ID P91143 PRELIMINARY; PRT; 372 AA.
 AC P91143;
 DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
 DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE SIMILAR TO ACETYLTTRANSFERASES.
 GN C37H5.2.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE: 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SUSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL NATURE 368:32-38(1994).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA DAVIDSON S., GILLAM B.,
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA WATSON R.,
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U88315; G1825777; -
 DR PFAM: PF00561; abhydrolase; 1.
 KW TRANSFERASE.
 SQ SEQUENCE 372 AA; 42139 MW; 5214FI59 CRC32;

Query Match 66.7%; Score 52; DB 5; Length 372;
 Best Local Similarity 58.3%; Pred. No. 1.08e+01;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 168 GGYLSTYALKY 179
 QY 1 ASYSTSSSLHY 12

RESULT 4
 ID 047582 PRELIMINARY; PRT; 531 AA.
 AC 047582;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE NADH DEHYDROGENASE SUBUNIT 5.
 OS ONCHOCERCA VOLVULUS.
 OG MITOCHONDRION.
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; SPIRURIA; SPIRURIDA;
 OC FILARIOIDEA; ONCHOCERCIDAE; ONCHOCERCA.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-FORREST;
 RA KEDDIE E.M., UNNASCH T.R.,
 RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF015193; G2735946; -
 KW MITOCHONDRION.
 SQ SEQUENCE 531 AA; 62891 MW; 47A734D9 CRC32;

Query Match 66.7%; Score 52; DB 8; Length 531;
 Best Local Similarity 50.0%; Pred. No. 1.08e+01;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 DB 268 FLAIGSLHY 277
 QY 3 YLSTSSSLHY 12

RESULT 5
 ID 061198 PRELIMINARY; PRT; 1238 AA.
 AC 061198;
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE F15E6.6 PROTEIN.
 GN F15E6.6.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE: 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SUSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL NATURE 368:32-38(1994).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA MILLER N., STELLIES L., BRADSHAW H., KEPLER D.,
 RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA WATSON R.,
 RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF038614; G2702437; -

DR PROSITE: PS00197; 2FEZS-FERREDOXIN; 1.
KW IRON-SULFUR.
SQ SEQUENCE 1238 AA; 135726 MW; 9408B7C CRC32;

Query Match
Best Local Similarity 70.0%; Score 52; DB 5; Length 1238;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 702 YLETOSSSLHY 711
11111111
QY 3 YLSTSSSLHY 12

RESULT 6
ID 077542 PRELIMINARY; PRT; 236 AA.
AC 077542;

DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
OLFACTORY RECEPTOR (FRAGMENT).

SUS SCROFA (PIG).
EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.

RN [1]
RP SEQUENCE FROM N.A.
RA MATARAZZO V., TIRARD A., RENUCCI M., BELAICH A., CLEMENT J.L.;
RT "Isolation of putative olfactory receptor sequences from pig nasal
epithelium";
RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF042016; G3273633; -.

FT NON_TER 1 1
FT NON_TER 236 236
SQ SEQUENCE 236 AA; 25564 MW; B9651DA1 CRC32;

Query Match
Best Local Similarity 50.0%; Score 51; DB 6; Length 236;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 66 YMAICSPHY 75
11111111
QY 3 YLSTSSSLHY 12

RESULT 7
ID 077543 PRELIMINARY; PRT; 238 AA.
AC 077543;

DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
OLFACTORY RECEPTOR (FRAGMENT).

SUS SCROFA (PIG).
EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.

RN [1]
RP SEQUENCE FROM N.A.
RA MATARAZZO V., TIRARD A., RENUCCI M., BELAICH A., CLEMENT J.L.;
RT "Isolation of putative olfactory receptor sequences from pig nasal
epithelium";
RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF042017; G3273635; -.

FT NON_TER 1 1
FT NON_TER 238 238
SQ SEQUENCE 238 AA; 25822 MW; BEF1C7F3 CRC32;

Query Match
Best Local Similarity 50.0%; Score 51; DB 6; Length 238;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 67 YMAICSPHY 76
11111111
QY 3 YLSTSSSLHY 12

RESULT 8
ID 070265 PRELIMINARY; PRT; 321 AA.
AC 070265;

DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE OLFACTORY RECEPTOR-LIKE PROTEIN.
GN SCR D-8.

OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; RODENTIA;
OC SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE; 98211958.

RA WALENSKY L.D., RUAT M., BAKIN R.E., BLACKSHAW S., RONNETT G.V.,
RT "Two novel odorant receptor families expressed in spermatids undergo
5'-splicing";
RL J. BIOL. CHEM. 273:9378-9387(1998).

DR EMBL; AF034897; G3153221; -.
SQ SEQUENCE 321 AA; 35375 MW; D85F971F CRC32;

Query Match
Best Local Similarity 50.0%; Score 51; DB 11; Length 321;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 129 YMAICSPHY 138
11111111
QY 3 YLSTSSSLHY 12

RESULT 9
ID 070267 PRELIMINARY; PRT; 321 AA.
AC 070267;

DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE OLFACTORY RECEPTOR-LIKE PROTEIN.

GN SCR D-9.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; RODENTIA;
OC SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE; 98211958.

RA WALENSKY L.D., RUAT M., BAKIN R.E., BLACKSHAW S., RONNETT G.V.,
RT "Two novel odorant receptor families expressed in spermatids undergo
5'-splicing";
RL J. BIOL. CHEM. 273:9378-9387(1998).

DR EMBL; AF034899; G3153225; -.
SQ SEQUENCE 321 AA; 35510 MW; D060B840 CRC32;

Query Match
Best Local Similarity 50.0%; Score 51; DB 11; Length 321;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 129 YMAICSPHY 138
11111111
QY 3 YLSTSSSLHY 12

RESULT 10
ID 088902 PRELIMINARY; PRT; 1494 AA.
AC 088902;

DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PROTEIN TYROSINE PHOSPHATASE TD14 (EC 3.1.3.48).
GN PTP-TD14.
OS RATTUS NORVEGICUS (RAT).

CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
 OC SCURIONATHI; MORIDAE; MORINAE; RATIUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN:
 RX MEDLINE; 98361981.
 RA CHOU L., ZHANG L., RUIZ-LOZANO P., YANG Q., CHIEN K.R., GRAHAM R.M.,
 RA ZHOU M.;
 RT "A novel putative protein-tyrosine phosphatase contains a BRO1-like
 RT domain and suppresses Ha-ras-mediated transformation.";
 RJ J. Biol. Chem. 273:21077-21083(1998).
 DR EMBL; AF077000; G3598974; -;
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 KW HYDROLASE.
 SQ SEQUENCE 1494 AA; 162931 MW; 9428F8BA CRC32;
 Query Match 65.4%; Score 51; DB 11; Length 1494;
 Best Local Similarity 36.4%; Pred. No. 1.68e+01;
 Matches 4; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
 QY 584 PFLGTAPLHF 594
 :|:|:|:|:|:
 2 SYLSTSSSLHY 12
 RESULT 11
 ID Q23863 PRELIMINARY; PRT; 2150 AA.
 AC Q23863;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HISTIDINE KINASE A.
 GN DHKA.
 OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
 CC EUKARYOTA; DICTYOSTELIIDA; DICTYOSTELIUM.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AM4;
 RX MEDLINE; 96324397.
 RA WANG N., SHAULISKY G., ESCALANTE R., LOOMIS W.F.;
 RT "A two-component histidine kinase gene that functions in
 RT Dictyostelium development.";
 RL EMO J. 15:3890-3896(1996).
 DR EMBL; U42597; G1136289; -;
 DR PFAM; PF00072; response_reg; 1.
 DR PFAM; PF00512; signal; 1.
 SQ SEQUENCE 2150 AA; 239662 MW; 02F946CE CRC32;
 Query Match 65.4%; Score 51; DB 5; Length 2150;
 Best Local Similarity 70.0%; Pred. No. 1.68e+01;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Db 180 NYLNSSSLH 189
 :|:|:|:|:|:
 2 SYLSTSSSLH 11
 RESULT 12
 ID Q23388 PRELIMINARY; PRT; 2219 AA.
 AC Q23388;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
 DE ZK1067.2 PROTEIN.
 GN ZK1067.2
 OS CAENORHABDITIS ELEGANS.
 CC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA THOMAS K.;
 RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE; 94150718.
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BOWFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HARKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SUSTON J.,
 RA THERREY-WIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDRAN P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL NATURE 368:32-38(1994).
 DR EMBL; 270038; E1351037; -;
 SQ SEQUENCE 2219 AA; 253649 MW; 59DEB843 CRC32;
 Query Match 65.4%; Score 51; DB 5; Length 2219;
 Best Local Similarity 50.0%; Pred. No. 1.68e+01;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 Db 628 SSYMRNGSLHF 639
 :|:|:|:|:|:
 1 ASYLSTSSSLHY 12
 RESULT 13
 ID P70013 PRELIMINARY; PRT; 220 AA.
 AC P70013;
 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
 DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE OLFACTORY RECEPTOR (FRAGMENT).
 OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
 OC MESOBATRACHIA; PIPOIDEA; PIPIIDAE; XENODODINAE; XENOPUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA FREITING J.;
 RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96112032.
 RA FREITING J., KRIGER J., STROTMAN J., BREER H.;
 RT "Two classes of olfactory receptors in xenopus laevis.";
 RL NEURON 15:1383-1392(1995).
 DR EMBL; Y08351; E273932; -;
 DR PFAM; PF00001; 7tm_1; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 220 AA; 24765 MW; EBD97044 CRC32;
 Query Match 64.1%; Score 50; DB 13; Length 220;
 Best Local Similarity 50.0%; Pred. No. 2.58e+01;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Db 60 YLACKPLHY 69
 :|:|:|:|:|:
 3 YLSTSSSLHY 12
 RESULT 14
 ID O45973 PRELIMINARY; PRT; 357 AA.
 AC O45973;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
 DE Y6E2A.6 PROTEIN.
 GN Y6E2A.6
 OS CAENORHABDITIS ELEGANS.
 CC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.

RN [1]
 RP SEQUENCE FROM N.A.
 RA MATTHEWS L.;
 RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HARKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALLDON N., SMITH A., SONNHAMMER E., SPADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
 RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.*;
 RT NATURE 368:32-38(1994).
 EMBL: AL021175; E1350810; -
 SEQUENCE 357 AA: 40700 MW: 985A9A9E CRC32:

Query Match 64.1%; Score 50; DB 5; Length 357;
 Best Local Similarity 50.0%; Pred. No. 2.58e+01;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 261 YLNGSAFHY 270
 Y 11:11:11
 QY 3 YLSTSSSLHY 12

RESULT 15
 ID 082449 PRELIMINARY; PRT; 419 AA.
 AC 082449;
 DT 01-NOV-1998 (TREMBLEREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLEREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLEREL. 08, LAST ANNOTATION UPDATE)
 DE P-HYDROXYPHENYLPYRUVATE DIOXYGENASE.
 GN PDS1
 OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
 OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
 OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
 OC CAPRIFALES; BRASSICACEAE; ARABIDOPSIS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV.COLUMBIA;
 RX MEDLINE: 98369110.
 NORRIS S.R., SHEN X., DELLA PENNA D.;
 RT *Complementation of the arabidopsis pds1 mutation with the gene
 RT encoding p-hydroxyphenylpyruvate dioxygenase.*;
 RL PLANT PHYSIOL. 117:1317-1323(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA DELLAPENNA D., NORRIS S.R., SHEN X.;
 RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF060481; G3694811; -
 KW DIOXYGENASE; PYRUVATE.
 SQ SEQUENCE 419 AA: 45854 MW: 60D99C92 CRC32:

Query Match 64.1%; Score 50; DB 10; Length 419;
 Best Local Similarity 58.3%; Pred. No. 2.58e+01;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 86 ASYLITSGDLRF 97
 Y 11:11:11
 QY 1 ASYLSTSSSLHY 12

Search completed: Thu Sep 2 12:44:04 1999
 Job time : 28 secs.

THIS PAGE BLANK (USPTO)

 W E S E R (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm
 on: Thu Sep 2 12:46:53 1999; Maspar time 3.13 Seconds
 Tabular output not generated. 153.392 Million cell updates/sec

Title: >US-08-599-226-33
 Description: (1-12) from US08599226.pep
 Perfect Score: 74
 Sequence: 1 ASFLSTSSSLRY 12

Scoring table: PAM 150
 Gap 15

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: p1r60
 1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 24.566; Variance 29.900; scale 0.822

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	54	73.0	469	2	D70048	ABC transporter (amln
2	54	73.0	1069	2	T00040	BH-protocadherin PCDH
3	54	73.0	1069	2	T00043	BH-protocadherin-a -
4	54	73.0	1072	2	T00041	BH-protocadherin PCDH
5	54	73.0	1200	2	T00042	BH-protocadherin PCDH
6	53	71.6	112	4	S59333	hypothetical protein
7	52	70.3	517	2	D37831	phenol 2-monooxygenas
8	52	70.3	949	3	T03030	hypothetical protein
9	50	67.6	296	2	S45336	finger protein, Spi/e
10	50	67.6	311	2	S13808	protein-tyrosine kina
11	50	67.6	346	2	S13809	protein-tyrosine kina
12	50	67.6	346	2	S13807	protein-tyrosine kina
13	50	67.6	464	2	B64970	hypothetical protein
14	50	67.6	471	2	S01052	flavonol 3-O-glucosyl
15	50	67.6	471	2	S08325	flavonol 3-O-glucosyl
16	50	67.6	471	2	S01037	flavonol 3-O-glucosyl
17	50	67.6	503	2	F64713	protein-export membra
18	50	67.6	511	2	S47290	phenol 2-monooxygenas
19	50	67.6	526	2	D71805	protein-export membra
20	50	67.6	529	2	UC5535	scavenger receptor Cl
21	50	67.6	593	2	S49525	glycoprotein G - siml
22	50	67.6	662	2	T01533	hypothetical protein
23	50	67.6	1166	2	S06142	kinase-related transf

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
24	49	66.2	155	2	B70169	hypothetical protein
25	49	66.2	339	2	S73840	ribonucleotide reduct
26	49	66.2	381	2	G71906	probable transcriptio
27	49	66.2	439	1	KHO8RT	cysteine proteinase (
28	49	66.2	509	1	A53920	scavenger receptor SR
29	49	66.2	516	2	S44306	phenol hydroxylase -
30	49	66.2	530	2	S62439	hypothetical protein
31	48	64.9	500	2	S49302	AMU1218 protein - whe
32	48	64.9	781	2	S43534	integrin betas - chic
33	48	64.9	1495	2	S61023	hypothetical protein
34	47	63.5	229	2	S33182	probable transport pr
35	47	63.5	238	1	Q0BE74	US4 protein - human h
36	47	63.5	262	2	A53452	phenylalanine hydroxy
37	47	63.5	265	2	S38380	Hrox1 protein - Calif
38	47	63.5	287	2	S57770	xyloglucan endo-1,4-b
39	47	63.5	342	2	S63654	hypothetical protein
40	47	63.5	457	2	E70081	putrine-cytosine perme
41	47	63.5	509	1	A48528	membrane glycoprotein
42	47	63.5	568	1	A46339	hemagglutinin precurs
43	47	63.5	668	2	H71312	probable ATP-dependen
44	47	63.5	792	2	S03232	hypothetical protein
45	47	63.5	828	2	S34695	hypothetical protein

ALIGNMENTS

RESULT	1	ALIGNMENTS
ENTRY	D70048	#type complete
TITLE	ABC transporter (amino acid permease) homolog yvsh - Bacillus subtilis	
ORGANISM	#formal_name Bacillus subtilis	
DATE	05-Dec-1997	#sequence_revision 05-Dec-1997
	24-Sep-1998	#text_change

REFERENCES

ACCESSIONS
 D70048
 A69580
 #authors

Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Boris, R.; Boursier, L.; Brans, A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.; Conerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devigne, K.M.; Duesterhoft, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.; Guisepi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.; Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi, Y.; Koeltter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, V.; Pohl, T.M.; Portetelle, D.; Porwolik, S.; Prescott, V.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenbol, M.; Vannier, F.; Vassalotti, A.; Viari, A.; Mambuti, R.; Wedler, E.; Wedler, H.; Weitzenecker, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

Nature (1997) 390:249-256
 The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

#cross-references MU1D:98044033
#accession D70048
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-469 ##label KUN
#cross-references GB:299121; GB:AL009126; NID:g2635827; PID:el166022;
PID:g2635847
#experimental_source strain 168
GENETICS
#gene yvsh
CLASSIFICATION #superfamily L-lysine transport protein
SUMMARY #length 469 #molecular-weight 50258 #checksum 4200
Query Match 73.0%; Score 54; DB 2; Length 469;
Best Local Similarity 45.5%; Pred. No. 2.14e+00;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Db 376 TFLTTAATLAY 386
QY 2 SFLSTSSLEY 12
RESULT 2
ENTRY T00040 #type complete
TITLE BH-protocadherin PCDH7 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change
ACCESSION T00040
REFERENCE 214074
#authors Yoshida, K.; Yoshitomo-Nakagawa, K.; Seki, N.; Sasaki, M.;
Sugano, S.
#journal Genomics (1998) 49:458-461
#title Cloning, expression analysis, and chromosomal localization of
BH-protocadherin (PCDH7), a novel member of the cadherin
superfamily.
#accession T00040
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-1069 ##label YOS
#cross-references EMBL:AB06755; NID:d1184677; PID:d1026122
#experimental_source clone BH-Pcdh-a
GENETICS
#map_position 4p15
SUMMARY #length 1069 #molecular-weight 116104 #checksum 9974
Query Match 73.0%; Score 54; DB 2; Length 1069;
Best Local Similarity 60.0%; Pred. No. 2.14e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 486 FLHTSTPIDY 495
QY 3 FLSTSSLEY 12
RESULT 3
ENTRY T00043 #type complete
TITLE BH-protocadherin-a - mouse
ORGANISM #formal_name mus musculus #common_name house mouse
DATE 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change
ACCESSION T00043
REFERENCE 214075
#authors Yoshida, K.
#submission submitted to the EMBL Data Library, August 1997
#accession T00043
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-1069 ##label YOS
#cross-references EMBL:AB06758; NID:d1227200; PID:d1033562
GENETICS
#gene pcdh7

#map_position 5C3-D
SUMMARY #length 1069 #molecular-weight 116313 #checksum 4821
Query Match 73.0%; Score 54; DB 2; Length 1069;
Best Local Similarity 60.0%; Pred. No. 2.14e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 486 FLHTSAPIDY 495
QY 3 FLSTSSLEY 12
RESULT 4
ENTRY T00041 #type complete
TITLE BH-protocadherin PCDH7 (clone BH-Pcdh-b) - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change
ACCESSION T00041
REFERENCE 214074
#authors Yoshida, K.; Yoshitomo-Nakagawa, K.; Seki, N.; Sasaki, M.;
Sugano, S.
#journal Genomics (1998) 49:458-461
#title Cloning, expression analysis, and chromosomal localization of
BH-protocadherin (PCDH7), a novel member of the cadherin
superfamily.
#accession T00041
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-1072 ##label YOS
#cross-references EMBL:AB06756; NID:d1184678; PID:d1026123
#experimental_source clone BH-Pcdh-b
GENETICS
#map_position 4p15
SUMMARY #length 1072 #molecular-weight 116462 #checksum 9727
Query Match 73.0%; Score 54; DB 2; Length 1072;
Best Local Similarity 60.0%; Pred. No. 2.14e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 486 FLHTSTPIDY 495
QY 3 FLSTSSLEY 12
RESULT 5
ENTRY T00042 #type complete
TITLE BH-protocadherin PCDH7 (clone BH-Pcdh-c) - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change
ACCESSION T00042
REFERENCE 214074
#authors Yoshida, K.; Yoshitomo-Nakagawa, K.; Seki, N.; Sasaki, M.;
Sugano, S.
#journal Genomics (1998) 49:458-461
#title Cloning, expression analysis, and chromosomal localization of
BH-protocadherin (PCDH7), a novel member of the cadherin
superfamily.
#accession T00042
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-1200 ##label YOS
#cross-references EMBL:AB06757; NID:d1184679; PID:d1026124
#experimental_source clone BH-Pcdh-c
GENETICS
#map_position 4p15
SUMMARY #length 1200 #molecular-weight 130337 #checksum 7152
Query Match 73.0%; Score 54; DB 2; Length 1200;
Best Local Similarity 60.0%; Pred. No. 2.14e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 439 FLHTSTPLDY 448
|||:|||||
QY 3 FLSTSSLEY 12

RESULT 6
ENTRY S59333 #type complete
TITLE hypothetical protein U3155 (U3149 internal orf) - yeast
ORGANISM Saccharomyces cerevisiae
#formal_name Saccharomyces cerevisiae
DATE 29-Nov-1995 #sequence_revision 05-Sep-1996 #text_change 05-Sep-1996
ACCESSIONS S59333
REFERENCE S59313
#authors Delius, H.
#submission submitted to the EMBL Data Library, June 1995
#description 36.8 kb of S. cerevisiae chromosome XII including ACE2, CK11, PDC5, SLT1, PUT1 and tRNA-Asp.
#accession S59333
#status Conceptual translation of pseudogene
#molecule_type DNA
#residues 1-112 #label DEL
#cross-references EMBL:X91256; NID:9995686; PID:9995707
#experimental_source strain S288C
COMMENT There is no evidence that this sequence is expressed.
GENETICS
#map_position 12R
KEYWORDS pseudogene
SUMMARY #length 112 #molecular-weight 13520 #checksum 5962

Query Match 71.6%; Score 53; DB 4; Length 112;
Best Local Similarity 70.0%; Pred. No. 3.44e+00;
Matches 7: Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 21 FLTSSSLAY 30
|||:|||||
QY 3 FLSTSSLEY 12

RESULT 7
ENTRY D37831 #type complete
TITLE phenol 2-monooxygenase (EC 1.14.13.7) chain P3 - Pseudomonas
ORGANISM Pseudomonas sp.
#formal_name Pseudomonas sp.
DATE 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 31-Oct-1997
ACCESSIONS D37831
REFERENCE A37831
#authors Nordlund, I.; Powlowski, J.; Shingler, V.
#journal J. Bacteriol. (1990) 172:6826-6833
#title Complete nucleotide sequence and polypeptide analysis of multicompnent phenol hydroxylase from Pseudomonas sp. strain CF600.
#cross-references M01D:91072230
#accession D37831
#status Preliminary
#molecule_type DNA
#residues 1-517 #label NOR
#cross-references GB:M60276; GB:M37764; NID:g151449; PID:g151453
KEYWORDS oxidoreductase
SUMMARY #length 517 #molecular-weight 60522 #checksum 6342

Query Match 70.3%; Score 52; DB 2; Length 517;
Best Local Similarity 60.0%; Pred. No. 5.50e+00;
Matches 6: Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 101 FLTAVSPLEY 110
|||:|||||
QY 3 FLSTSSLEY 12
RESULT 8
ENTRY T03030 #type fragment

TITLE hypothetical protein KIA00365 - human (fragment)
ORGANISM Homo sapiens
#formal_name Homo sapiens
#common_name man
DATE 23-Mar-1999 #sequence_revision 23-Mar-1999 #text_change 23-Mar-1999
ACCESSIONS T03030
REFERENCE Z14651

#authors Lamerdin, J.E.; McCready, P.M.; Skowronski, E.; Adamson, A.W.; Burkhardt-Schultz, K.; Gordon, L.; Kyle, A.; Ramirez, M.; Stillwagen, S.; Phan, H.; Velasco, N.; Ganes, J.; Dangann, L.; Poundstone, P.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Altix, C.; Andreise, T.; Tranheim, M.; Amico-Keller, G.; Coefield, J.; Duarte, S.; Lucas, S.; Bruce, R.; Thomas, P.; Quan, G.; Kronmiller, B.; Arellano, A.; Montgomery, M.; Ow, D.; Nolan, M.
#submission submitted to the EMBL Data Library, March 1998
#description Sequence analysis of an ~1 Mb region containing the MEF2B gene in 19p12.
#accession T03030
#status Preliminary
#residues 1-949 #label LAM
#cross-references EMBL:AC004447; NID:g2978446; PID:g2978447
SUMMARY #length 949 #checksum 5364

Query Match 70.3%; Score 52; DB 3; Length 949;
Best Local Similarity 70.0%; Pred. No. 5.50e+00;
Matches 7: Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 475 FLSDNSLEY 484
|||:|||||
QY 3 FLSTSSLEY 12

RESULT 9
ENTRY S45336 #type complete
TITLE finger protein, Spi/egr-like - fruit fly (Drosophila sp.)
ORGANISM Drosophila sp.
#formal_name Drosophila sp.
DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997
ACCESSIONS S45336
REFERENCE S45336
#authors Broenner, G.; Chu-Lagriff, O.; Doe, C.O.; Cohen, B.; Weigel, D.; Taubert, H.; Jaekle, H.
#journal Nature (1994) 369:664-668
#title Spi/egr-like zinc-finger protein required for endoderm specification and germ-layer formation in Drosophila.
#accession S45336
#status Preliminary
#molecule_type DNA
#residues 1-296 #label BRO
SUMMARY #length 296 #molecular-weight 33598 #checksum 2521

Query Match 67.6%; Score 50; DB 2; Length 296;
Best Local Similarity 58.3%; Pred. No. 1.38e+01;
Matches 7: Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Db 127 APFLASDLYX 138
|||:|||||
QY 1 ASPLSTSSLEY 12

RESULT 10
ENTRY S13808 #type fragment
TITLE protein-tyrosine kinase (EC 2.7.1.112) mrk (X chromosome) - southern platyfish (fragment)
ORGANISM melanoma-inducing protein
#formal_name Xiphophorus maculatus #common_name southern platyfish
DATE 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 30-Jan-1998
ACCESSIONS S13808
REFERENCE S13807
#authors Adam, D.; Maeuener, W.; Scharlt, M.
#journal Oncogene (1991) 6:73-80

#title Transcriptional activation of the melanoma inducing Xmrk
#cross-references EMBL:X56318; NID:g65282; PID:g65283
#accession S13807
#status translation not shown
#molecule-type DNA
#residues 1-311 ##label ADA
#cross-references EMBL:X56318; NID:g65282; PID:g65283

GENETICS
#map.position X
#introns 52/3; 78/1; 127/1; 159/3; 205/3
CLASSIFICATION #superfamily epidermal growth factor receptor; protein kinase
#homology
KEYWORDS ATP; phosphotransferase
FEATURE 1-152
SUMMARY #length 311 #checksum 2822

Query Match 67.6%; Score 50; DB 2; Length 311;
Best Local Similarity 50.0%; Pred. No. 1.38e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 290 FLPAENLEY 299
|||:::||||
QY 3 FLSTSSLEY 12

RESULT 11
ENTRY S13807 #type fragment
TITLE protein-tyrosine kinase (EC 2.7.1.112) mrk (Y chromosome) -
ALTERNATE_NAMES southern platyfish (fragment)
ORGANISM melanoma-inducing protein
#formal_name Xiphophorus maculatus #common_name southern
platyfish
DATE 21-Nov-1993 #sequence_revision 13-Mar-1997 #text_change
08-Sep-1997

ACCESSIONS S13809
REFERENCE S13807
#authors Adam, D.; Maeucler, W.; Scharl, M.
#journal Oncogene (1991) 6:73-80
#title Transcriptional activation of the melanoma inducing Xmrk
oncogene in Xiphophorus.
#cross-references MUID:91125882
#accession S13809
#status preliminary; translation not shown
#molecule-type DNA
#residues 1-346 ##label ADA
#cross-references EMBL:X56319; NID:g65284; PID:g65285

GENETICS
#map.position Y
#introns 52/3; 78/1; 127/1; 159/3; 205/3; 236/1
CLASSIFICATION #superfamily epidermal growth factor receptor; protein kinase
#homology
KEYWORDS phosphotransferase
FEATURE 1-152
SUMMARY #length 346 #checksum 797

Query Match 67.6%; Score 50; DB 2; Length 346;
Best Local Similarity 50.0%; Pred. No. 1.38e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 325 FLPAENLEY 334
|||:::||||
QY 3 FLSTSSLEY 12

RESULT 12
ENTRY S13807 #type fragment
TITLE protein-tyrosine kinase (EC 2.7.1.112) mrk (invariably
present) - southern platyfish (fragment)
ALTERNATE_NAMES melanoma-inducing protein
ORGANISM #formal_name Xiphophorus maculatus #common_name southern

DATE platyfish
21-Nov-1993 #sequence_revision 30-Jan-1998 #text_change
30-Jan-1998

ACCESSIONS S13807
REFERENCE S13807
#authors Adam, D.; Maeucler, W.; Scharl, M.
#journal Oncogene (1991) 6:73-80
#title Transcriptional activation of the melanoma inducing Xmrk
oncogene in Xiphophorus.
#cross-references EMBL:X56317; NID:g65280; PID:g65281

GENETICS
#accession S13807
#status translation not shown
#molecule-type DNA
#residues 1-346 ##label ADA
#cross-references EMBL:X56317; NID:g65280; PID:g65281

KEYWORDS mrk
FEATURE 52/3; 78/1; 127/1; 159/3; 205/3; 236/1
SUMMARY #length 346 #checksum 1508

Query Match 67.6%; Score 50; DB 2; Length 346;
Best Local Similarity 50.0%; Pred. No. 1.38e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 325 FLPAENLEY 334
|||:::||||
QY 3 FLSTSSLEY 12

RESULT 13
ENTRY B64970 #type complete
TITLE hypothetical protein b2043 - Escherichia coli (strain K-12)
ORGANISM #formal_name Escherichia coli
DATE 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change
14-Nov-1997

ACCESSIONS B64970
REFERENCE A64720
#authors Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.;
Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.;
Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.;
Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao,
Y.
#journal Science (1997) 277:1453-1462
#title The complete genome sequence of Escherichia coli K-12.
#cross-references MUID:97426617
#accession B64970
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule-type DNA
#residues 1-464 ##label BLAT
#cross-references GB:AB000295; GB:U00096; NID:g1788354; PID:g1788356;
UWGP:b2043

SUMMARY #experimental source strain K-12, substrain MG1655
#length 464 #molecular-weight 51315 #checksum 8372

Query Match 67.6%; Score 50; DB 2; Length 464;
Best Local Similarity 45.5%; Pred. No. 1.38e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 10 TFLASSALAF 20
:|:::|::|:
QY 2 FLSTSSLEY 12

RESULT 14
ENTRY S01052 #type complete
TITLE flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (allele
Bz-MC) - maize

ALTERNATE_NAMES UDPglucose flavonoid glucosyltransferase
ORGANISM #formal_name Zea mays #common_name maize
DATE 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change
08-Sep-1997
ACCESSIONS S01052; S08324
REFERENCE S01037
#authors Raistson, E.J.; English, J.J.; Dooner, H.K.
#journal Genetics (1988) 119:185-197
#title Sequence of three bronze alleles of maize and correlation
with the genetic fine structure.
#cross-references MUID:88284304
#accession S01052
#status translation not shown
#molecule_type DNA
#residues 1-471 #label RAL
REFERENCE #cross-references EMBL:X07940; NID:g22204; PID:g22205
S08324
#authors Furtak, D.; Schiefelbein, J.W.; Johnston, F.; Nelson Jr.,
O.E.
#journal Plant Mol. Biol. (1988) 11:473-481
#title Sequence comparisons of three wild-type Bronze-1 alleles from
Zea mays.
#accession S08324
#status translation not shown
#molecule_type DNA
#residues 1-471 #label FUR
#cross-references EMBL:X13500; NID:g22364; PID:g1030071
GENETICS
#gene Bz1
#map_position 9
#introns 175/1
CLASSIFICATION #superfamily flavonol O3-glucosyltransferase
KEYWORDS glycosyltransferase; hexosyltransferase
SUMMARY #length 471 #molecular-weight 48769 #checksum 6660
Query Match 67.6%; Score 50; DB 2; Length 471;
Best Local Similarity 88.9%; Pred. No. 1.38e+01;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
DB 47 SFLSTASSL 55
QY 2 SFLSTSSSL 10
RESULT 15
ENTRY S08325 #type complete
#title Flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (allele
B2MCC2) - maize
ALTERNATE_NAMES UDPglucose flavonoid glucosyl-transferase
ORGANISM #formal_name Zea mays #common_name maize
DATE 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change
08-Sep-1997
ACCESSIONS S08325
REFERENCE S08324
#authors Furtak, D.; Schiefelbein, J.W.; Johnston, F.; Nelson Jr.,
O.E.
#journal Plant Mol. Biol. (1988) 11:473-481
#title Sequence comparisons of three wild-type Bronze-1 alleles from
Zea mays.
#accession S08325
#status translation not shown
#molecule_type DNA
#residues 1-471 #label FUR
#cross-references EMBL:X13501; NID:g22361; PID:g295854
GENETICS
#gene Bz1
#introns 175/1
CLASSIFICATION #superfamily flavonol O3-glucosyltransferase
KEYWORDS glycosyltransferase; hexosyltransferase
SUMMARY #length 471 #molecular-weight 48621 #checksum 7439
Query Match 67.6%; Score 50; DB 2; Length 471;
Best Local Similarity 88.9%; Pred. No. 1.38e+01;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
DB 47 SFLSTASSL 55
QY 2 SFLSTSSSL 10

Search completed: Thu Sep 2 12:47:12 1999
Job time : 19 secs.

THIS PAGE BLANK (USPTO)

MUSEUM (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSrch_pp protein - protein database search, using Smith-Waterman algorithm
on: Thu Sep 2 12:48:34 1999; Maspar time 1.45 Seconds
83.761 Million cell updates/sec
Tabular output not generated.

Title: >US-08-599-226-33
Description: (1-12) from US08599226.pep
Perfect Score: 74
Sequence: 1 ASFLSTSSSLLEY 12

Scoring table:
PAM 150
Gap 15

Searched: 106580 seqs, 10152877 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A.COMB 2:5B.COMB 3:PCN.COMB 4:backfiles1

Statistics: Mean 16.342; Variance 51.433; scale 0.318

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	52	70.3	1026	3	PCT-US95-0 Sequence 95, Applicati	3.63e+01
2	52	70.3	1026	1	US-07-998 Sequence 95, Applicati	3.63e+01
3	52	70.3	1026	1	US-08-453 Sequence 95, Applicati	3.63e+01
4	52	70.3	1026	1	US-08-453 Sequence 95, Applicati	3.63e+01
5	52	70.3	1026	3	PCT-US93-1 Sequence 95, Applicati	3.63e+01
6	52	70.3	1026	2	US-08-268 Sequence 95, Applicati	3.63e+01
7	52	70.3	1026	2	PCT-US95-0 Sequence 103, Applicat	3.63e+01
8	52	70.3	1203	1	US-07-998 Sequence 103, Applicat	3.63e+01
9	52	70.3	1203	1	US-08-453 Sequence 103, Applicat	3.63e+01
10	52	70.3	1203	1	US-08-453 Sequence 103, Applicat	3.63e+01
11	52	70.3	1203	3	PCT-US93-1 Sequence 103, Applicat	3.63e+01
12	52	70.3	1203	2	US-08-268 Sequence 103, Applicat	3.63e+01
13	50	67.6	223	2	US-08-985 Sequence 11, Applicati	5.88e+01
14	50	67.6	445	2	US-08-985 Sequence 2, Applicati	5.88e+01
15	50	67.6	509	3	PCT-US95-0 Sequence 8, Applicati	7.47e+01
16	49	66.2	509	3	PCT-US95-0 Sequence 4, Applicati	9.47e+01
17	48	64.9	158	2	US-08-828 Sequence 3, Applicati	9.47e+01
18	48	64.9	159	2	US-08-828 Sequence 1, Applicati	1.20e+02
19	47	63.5	223	2	US-08-985 Sequence 17, Applicati	1.20e+02
20	47	63.5	362	2	US-08-985 Sequence 4, Applicati	1.52e+02
21	46	62.2	218	2	US-08-336 Sequence 4, Applicati	1.52e+02
22	46	62.2	238	2	US-08-928 Sequence 5, Applicati	1.52e+02
23	46	62.2	238	2	US-08-928 Sequence 5, Applicati	1.52e+02

RESULT	1	STANDARD	PRT	1026	AA
ID	PCT-US95-08071-95				
XX	xxxxxx				
XX	Sequence 95, Application PC/TUS9508071				
DE	Sequence 95, Application PC/TUS9508071				
CC	GENERAL INFORMATION:				
CC	APPLICANT: Suzuki, Shintaro				
CC	TITLE OF INVENTION: Protocadherin Materials and Methods				
CC	NUMBER OF SEQUENCES: 115				
CC	CORRESPONDENCE ADDRESSES:				
CC	ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &				
CC	ADDRESSEE: Borun				
CC	STREET: 6300 Sears Tower, 233 S. Wacker Drive				
CC	CITY: Chicago				
CC	STATE: Illinois				
CC	COUNTRY: USA				
CC	ZIP: 60606				
CC	COMPUTER READABLE FORM:				
CC	MEDIUM TYPE: Floppy disk				
CC	COMPUTER: IBM PC compatible				
CC	OPERATING SYSTEM: PC-DOS/MS-DOS				
CC	SOFTWARE: Patent Release #1.0, Version #1.25				
CC	CURRENT APPLICATION DATA:				
CC	APPLICATION NUMBER: PCT/US95/08071				
CC	FILING DATE:				
CC	CLASSIFICATION:				
CC	Prior APPLICATION DATA:				
CC	APPLICATION NUMBER: PCT/US93/12588				
CC	FILING DATE: 23 DEC 1993				
CC	Prior APPLICATION DATA:				
CC	APPLICATION NUMBER: US 07/998, 003				
CC	FILING DATE: 29 DEC 1992				
CC	ATTORNEY/AGENT INFORMATION:				
CC	NAME: Noland, Greta E.				
CC	REGISTRATION NUMBER: 35,302				
CC	REFERENCE/DOCKET INFORMATION:				
CC	TELECOMMUNICATION INFORMATION:				
CC	TELEPHONE: 312/474-6300				
CC	TELEFAX: 312/474-0448				
CC	TELEX: 25-3856				
CC	INFORMATION FOR SEQ ID NO: 95:				
CC	SEQUENCE CHARACTERISTICS:				

CC LENGTH: 1026 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: Protein
SQ SEQUENCE 1026 AA; 111270 MW; 5611711 CN;

Query Match 70.3%; Score 52; DB 3; Length 1026;
Best Local Similarity 50.0%; Pred. No. 3.63e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 423 FLQTTTPLY 432
|||:::|:|
OY 3 FLSTSSSLEY 12

RESULT 2
ID US-07-998-003A-95 STANDARD; PRT: 1026 AA.
XX
AC xxxxxx

Sequence 95, Application US/07998003A

CC Sequence 95, Application US/07998003A
CC Patent No. 5643781
CC GENERAL INFORMATION:
CC APPLICANT: Suzuki, Shintaro
CC TITLE OF INVENTION: Protocadherin Materials and Methods
CC NUMBER OF SEQUENCES: 107
CC CORRESPONDENCE ADDRESS:
CC

CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
CC ADDRESSEE: Bicknell
CC STREET: 20 South Clark Street
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60603

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/998,003A
CC FILING DATE:
CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:
CC NAME: No. 5643781and, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 30903
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/346-5750
CC TELEFAX: 312/984-9740
CC TELEX: 25-3856

CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1026 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: Protein
SQ SEQUENCE 1026 AA; 111270 MW; 5611711 CN;

Query Match 70.3%; Score 52; DB 1; Length 1026;
Best Local Similarity 50.0%; Pred. No. 3.63e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 423 FLQTTTPLY 432
|||:::|:|
OY 3 FLSTSSSLEY 12

RESULT 3

ID US-08-453-695A-95 STANDARD; PRT: 1026 AA.
XX
AC xxxxxx

Sequence 95, Application US/08453695A
CC Sequence 95, Application US/08453695A
CC Patent No. 5708143

CC GENERAL INFORMATION:
CC APPLICANT: Suzuki, Shintaro
CC TITLE OF INVENTION: Protocadherin Materials and Methods
CC NUMBER OF SEQUENCES: 115
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
CC STREET: 233 South Wacker, 6300 Sears Tower
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/453,695A
CC FILING DATE:
CC CLASSIFICATION: 530

CC ATTORNEY/AGENT INFORMATION:
CC NAME: No. 5708143and, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 32658
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/474-6300
CC TELEFAX: 312/474-0448
CC TELEX: 25-3856

CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1026 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: Protein
SQ SEQUENCE 1026 AA; 111270 MW; 5611711 CN;

Query Match 70.3%; Score 52; DB 1; Length 1026;
Best Local Similarity 50.0%; Pred. No. 3.63e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 423 FLQTTTPLY 432
|||:::|:|
OY 3 FLSTSSSLEY 12

RESULT 4
ID US-08-453-274B-95 STANDARD; PRT: 1026 AA.
XX
AC xxxxxx

Sequence 95, Application US/08453274B

CC Sequence 95, Application US/08453274B
CC Patent No. 5663300

CC GENERAL INFORMATION:
CC APPLICANT: Suzuki, Shintaro
CC TITLE OF INVENTION: Protocadherin Materials and Methods
CC NUMBER OF SEQUENCES: 107
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

CC STREET: 6300 Sears Tower, 233 South Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: United States of America
CC ZIP: 60606-6402
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/453,274B
CC FILING DATE: 30-MAY-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: No. 5663300and, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 32860
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/474-6300
CC TELEFAX: 312/474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1026 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1026 AA; 111270 MW; 5611711 CN;
SQ
Query Match 70.3%; Score 52; DB 1; Length 1026;
Best Local Similarity 50.0%; Pred. No. 3.63e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 423 FLOTTPLDY 432
1111111111
0Y 3 FLSTSSLEY 12
RESULT 5
ID PCT-US93-12588-95 STANDARD; PRT; 1026 AA.
XX xxxxxx
AC
DT
XX
DE Sequence 95, Application PC/TUS9312588
CC
CC GENERAL INFORMATION:
CC APPLICANT: Suzuki, Shintaro
CC TITLE OF INVENTION: Protocadherin Materials and Methods
CC NUMBER OF SEQUENCES: 107
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
CC ADDRESSEE: Borun
CC STREET: 6300 Sears Tower, 233 S. Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/12588
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/998,003
CC FILING DATE: 29 DEC 1992
CC ATTORNEY/AGENT INFORMATION:

CC NAME: Noland, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 31811
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/474-6300
CC TELEFAX: 312/474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1026 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1026 AA; 111270 MW; 5611711 CN;
SQ
Query Match 70.3%; Score 52; DB 3; Length 1026;
Best Local Similarity 50.0%; Pred. No. 3.63e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 423 FLOTTPLDY 432
1111111111
0Y 3 FLSTSSLEY 12
RESULT 6
ID US-08-268-161A-95 STANDARD; PRT; 1026 AA.
XX xxxxxx
AC
DT
XX
XX
DE Sequence 95, Application US/08268161A
CC
CC Sequence 95, Application US/08268161A
CC Patent No. 5798224
CC GENERAL INFORMATION:
CC APPLICANT: Suzuki, Shintaro
CC TITLE OF INVENTION: Protocadherin Materials and Methods
CC NUMBER OF SEQUENCES: 115
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
CC ADDRESSEE: Borun
CC STREET: 233 South Wacker, 6300 Sears Tower
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/268,161A
CC FILING DATE: June 27, 1994
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Young J. Suh
CC REGISTRATION NUMBER: P-41,337
CC REFERENCE/DOCKET NUMBER: 27866/32149
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/474-6300
CC TELEFAX: 312/474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1026 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1026 AA; 111270 MW; 5611711 CN;
SQ
Query Match 70.3%; Score 52; DB 2; Length 1026;

Best Local Similarity 50.0%; Pred. No. 3.63e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 423 FLOTTPLDY 432
|||::|:|
QY 3 FLSTSSLEY 12

RESULT 7
ID PCT-US95-08071-103 STANDARD; PRT; 1203 AA.
AC xxxxxx
XX

DE Sequence 103, Application PC/TUS9508071

CC Sequence 103, Application PC/TUS9508071

CC APPLICANT: Suzuki, Shintaro

CC TITLE OF INVENTION: Protocadherin Materials and Methods

CC NUMBER OF SEQUENCES: 115

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &

CC STREET: 6300 Sears Tower, 233 S. Wacker Drive

CC CITY: Chicago

CC STATE: Illinois

CC COUNTRY: USA

CC ZIP: 60606

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US95/08071

CC FILING DATE:

CC CLASSIFICATION:

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: PCT/US93/12588

CC FILING DATE: 23 DEC 1993

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 07/998,003

CC FILING DATE: 29 DEC 1992

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Noland, Grete E.

CC REGISTRATION NUMBER: 35,302

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 312/474-6300

CC TELEFAX: 312/474-0448

CC TELEX: 25-3856

CC INFORMATION FOR SEQ ID NO: 103:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1203 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

CC SEQUENCE 1203 AA; 130280 MW; 7658453 CN;

Query Match 70.3%; Score 52; DB 3; Length 1203;

Best Local Similarity 50.0%; Pred. No. 3.63e+01;

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 423 FLOTTPLDY 432
|||::|:|

QY 3 FLSTSSLEY 12

RESULT 8
ID US-07-998-003A-103 STANDARD; PRT; 1203 AA.

AC xxxxxx

XX

DE

XX

XX

Sequence 103, Application US/07998003A

CC Patent No. 5643781

CC GENERAL INFORMATION:

CC APPLICANT: Suzuki, Shintaro

CC TITLE OF INVENTION: Protocadherin Materials and Methods

CC NUMBER OF SEQUENCES: 107

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &

CC STREET: 20 South Clark Street

CC CITY: Chicago

CC STATE: Illinois

CC COUNTRY: USA

CC ZIP: 60603

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/07/998,003A

CC FILING DATE:

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: No. 5643781and, Grete E.

CC REGISTRATION NUMBER: 35,302

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 312/346-5750

CC TELEFAX: 312/984-9740

CC TELEX: 25-3856

CC INFORMATION FOR SEQ ID NO: 103:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1203 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

CC SEQUENCE 1203 AA; 130280 MW; 7658453 CN;

Query Match 70.3%; Score 52; DB 1; Length 1203;

Best Local Similarity 50.0%; Pred. No. 3.63e+01;

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 423 FLOTTPLDY 432
|||::|:|

QY 3 FLSTSSLEY 12

RESULT 9
ID US-08-453-695A-103 STANDARD; PRT; 1203 AA.

AC xxxxxx

XX

XX

DE

XX

Sequence 103, Application US/08453695A

CC Patent No. 5708143

CC GENERAL INFORMATION:

CC APPLICANT: Suzuki, Shintaro

CC TITLE OF INVENTION: Protocadherin Materials and Methods

CC NUMBER OF SEQUENCES: 115

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &

CC STREET: 233 South Wacker, 6300 Sears Tower

CC	CITY:	Chicago
CC	STATE:	Illinois
CC	COUNTRY:	USA
CC	ZIP:	60606
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE:	Floppy disk
CC	COMPUTER:	IBM PC compatible
CC	OPERATING SYSTEM:	PC-DOS/MS-DOS
CC	SOFTWARE:	Patentin Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER:	US/08/453, 695A
CC	FILING DATE:	
CC	CLASSIFICATION:	530
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME:	No. 5708143and, Greta E.
CC	REGISTRATION NUMBER:	35,302
CC	REFERENCE/DOCKET NUMBER:	32658
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE:	312/474-6300
CC	TELEFAX:	312/474-0448
CC	TELEX:	25-3856
CC	INFORMATION FOR SEQ. ID NO:	103:
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH:	1203 amino acids
CC	TYPE:	amino acid
CC	TOPOLOGY:	linear
CC	MOLECULE TYPE:	protein
CC	SEQUENCE	1203 AA, 130280 MM, 7658453 CN;
Db	Query Match	70.3%; Score 52; DB 1; Length 1203;
	Best Local Similarly	50.0%; Pred. NO. 3.63e+01;
QY	Matches	5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
	3 FLTSSLEY 12	
Db	423 FLOTTPLDY 432	
XX	RESULT 10	
XX	ID US-08-453-274B-103	STANDARD; PRT; 1203 AA.
XX	xxxxxx	
XX	Sequence 103, Application US/08453274B	
DE	Sequence 103, Application US/08453274B	
CC	Patent No. 5663300	
CC	GENERAL INFORMATION:	
CC	APPLICANT:	Suzuki, Shintaro
CC	TITLE OF INVENTION:	Proteodherin Materials and Methods
CC	NUMBER OF SEQUENCES:	107
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE:	Marshall, O'Toole, Gerstein, Murray & Borun
CC	STREET:	6300 Sears Tower, 233 South Wacker Drive
CC	City:	Chicago
CC	STATE:	Illinois
CC	COUNTRY:	United States of America
CC	ZIP:	60606-6402
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE:	Floppy disk
CC	COMPUTER:	IBM PC compatible
CC	OPERATING SYSTEM:	PC-DOS/MS-DOS
CC	SOFTWARE:	Patentin Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER:	US/08/453, 274B
CC	FILING DATE:	30-MAY-1995
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME:	No. 5663300and, Greta E.
CC	REGISTRATION NUMBER:	35,302
CC	REFERENCE/DOCKET NUMBER:	32660
CC	TELECOMMUNICATION INFORMATION:	

CC	TELEPHONE:	312/474-6300
CC	TELEFAX:	312/474-0448
CC	TELEX:	25-3856
CC	INFORMATION FOR SEQ ID NO:	103:
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH:	1203 amino acids
CC	TYPE:	amino acid
CC	TOPOLOGY:	linear
CC	MOLECULE TYPE:	protein
SQ	SEQUENCE	1203 AA; 130280 MW; 7658453 CN;
Query Match		
	Best Local Similarity	70.3%; Score 52; DB 1; Length 1203; Pred. No. 3.63e+01;
Matches	5; Conservative	4; Mismatches 1; Indels 0; Gaps 0;
Db	423 FLOTTPLUDY 432	
OY	:::~::~	
	3 FLSTSSSLEY 12	
RESULT	11	
ID	PCT-US93-12588-103	STANDARD; PRT; 1203 AA.
XX	AC	xxxxxx
XX	DT	
DE	Sequence 103, Application PC/TUS9312588	
CC	Sequence 103, Application PC/TUS9312588	
CC	GENERAL INFORMATION:	
CC	APPLICANT:	Suzuki, Shintaro
CC	TITLE OF INVENTION:	Protocadherin Materials and Methods
CC	NUMBER OF SEQUENCES:	107
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE:	Marshall, O'Toole, Gerstein, Murray, &
CC	ADDRESSEE:	Born
CC	STREET:	6300 Sears Tower, 233 S. Wacker Drive
CC	CITY:	Chicago
CC	STATE:	Illinois
CC	COUNTRY:	USA
CC	ZIP:	60606
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE:	Floppy disk
CC	COMPUTER:	IBM PC compatible
CC	OPERATING SYSTEM:	PC-DOS/MS-DOS
CC	SOFTWARE:	PatentIn Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER:	PCT/US93/12588
CC	FILING DATE:	
CC	CLASSIFICATION:	
CC	PRIOR APPLICATION DATA:	
CC	APPLICATION NUMBER:	US 07/998,003
CC	FILING DATE:	29 DEC 1992
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME:	Noland, Greta E.
CC	REGISTRATION NUMBER:	35,302
CC	REFERENCE/DOCKET NUMBER:	31811
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE:	312/474-6300
CC	TELEFAX:	312/474-0448
CC	TELEX:	25-3856
CC	INFORMATION FOR SEQ ID NO:	103:
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH:	1203 amino acids
CC	TYPE:	amino acid
CC	TOPOLOGY:	linear
CC	MOLECULE TYPE:	protein
SQ	SEQUENCE	1203 AA; 130280 MW; 7658453 CN;
Query Match		
	Best Local Similarity	70.3%; Score 52; DB 3; Length 1203; Pred. No. 3.63e+01;
Matches	5; Conservative	4; Mismatches 1; Indels 0; Gaps 0;

Db 423 FLQTTPLDY 432
1111111111
QY 3 FLSTSSSLEY 12

RESULT 12
ID US-08-268-161A-103 STANDARD; PRT: 1203 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 103, Application US/08268161A
XX
CC Sequence 103, Application US/08268161A
CC Patent No. 5798224
CC GENERAL INFORMATION:
CC APPLICANT: Suzuki, Shintaro
CC TITLE OF INVENTION: Protoporphyrin Materials and Methods
CC NUMBER OF SEQUENCES: 115
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
CC ADDRESSEE: Borun
CC STREET: 233 South Wacker, 6300 Sears Tower
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/268,161A
CC FILING DATE: June 27, 1994
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Young J. Suh
CC REGISTRATION NUMBER: P-41,337
CC REFERENCE/DOCKET NUMBER: 27866/32149
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/474-6300
CC TELEFAX: 312/474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 103:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1203 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1203 AA; 130280 MW; 7658453 CN;

Query Match 70.3%; Score 52; DB 2; Length 1203;
Best Local Similarity 50.0%; Pred. No. 3.63e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 423 FLQTTPLDY 432
1111111111
QY 3 FLSTSSSLEY 12

RESULT 13
ID US-08-985-090-11 STANDARD; PRT: 23 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 11, Application US/08985090
XX
CC Sequence 11, Application US/08985090

CC Patent No. 5885893
CC GENERAL INFORMATION:
CC APPLICANT: Andrew D. J. Goodearl
CC TITLE OF INVENTION: MUSCARINIC RECEPTORS AND USES THEREFOR
CC NUMBER OF SEQUENCES: 28
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD, LLP
CC STREET: 28 State Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/985,090
CC FILING DATE:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jean M. Silver1
CC REGISTRATION NUMBER: 39,030
CC REFERENCE/DOCKET NUMBER: MNI-032
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)742-4214
CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 23 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 23 AA; 2719 MW; 3615 CN;

Query Match 67.6%; Score 50; DB 2; Length 23;
Best Local Similarity 60.0%; Pred. No. 5.88e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 2 FLTASTLEF 11
1111111111
QY 3 FLSTSSSLEY 12

RESULT 14
ID US-08-985-090-2 STANDARD; PRT: 445 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 2, Application US/08985090
XX
CC Sequence 2, Application US/08985090
CC Patent No. 5885893
CC GENERAL INFORMATION:
CC APPLICANT: Andrew D. J. Goodearl
CC TITLE OF INVENTION: MUSCARINIC RECEPTORS AND USES THEREFOR
CC NUMBER OF SEQUENCES: 28
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD, LLP
CC STREET: 28 State Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS

```
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/985_090  
CC FILING DATE:  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER:  
CC FILING DATE:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Jean M. Silverl  
CC REGISTRATION NUMBER: 39,030  
CC REFERENCE/DOCKET NUMBER: MN1-032  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (617)227-7400  
CC TELEFAX: (617)742-4214  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 445 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 445 AA; 48671 MW; 1062278 CN;  
  
Db 198 FLITASTLEF 207  
|||::|||  
QY 3 FLTSTSSLEY 12  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
RESULT 15  
ID PCT-US95-07721-8 STANDARD: PRT: 509 AA.  
XX xxxxxx  
DT  
Sequence 8, Application PC/TUS9507721  
DE  
CC Sequence 8, Application PC/TUS9507721  
CC GENERAL INFORMATION:  
CC APPLICANT: Massachusetts Institute of Technology  
CC TITLE OF INVENTION: Class BI and CI Scavenger Receptors  
CC NUMBER OF SEQUENCES: 8  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Patrea L. Pabst  
CC STREET: 2800 One Atlantic Center  
CC STREET: 1201 West Peachtree Street  
CC CITY: Atlanta  
CC STATE: Georgia  
CC COUNTRY: USA  
CC ZIP: 30309-3450  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: IBM PC compatible  
CC SOFTWARE: PC-DOS/MS-DOS  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/07721  
CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Pabst, Patrea L.  
CC REGISTRATION NUMBER: 31,284  
CC REFERENCE/DOCKET NUMBER: MIT6620  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (404) 873-8794  
CC TELEFAX: (404) 873-8795  
CC INFORMATION FOR SEQ ID NO: 8:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 509 amino acids  
CC TYPE: amino acid
```

```
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      FEATURE:
CC      NAME/KEY: misc. feature
CC      LOCATION: 1..509
CC      OTHER INFORMATION: /Function = "Amino acid sequence for
CC      OTHER INFORMATION: the murine scavenger Receptor Class BI."
SQ      SEQUENCE 509 AA; 56754 MW; 1451442 CN;

Query Match          67.6%; Score 50; DB 3; Length 509;
Best Local Similarity 50.0%; Pred. No. 5.88e+01;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db    263 APFMTPESSLEF 274
      I:::|::|::
QY    1 ASFLSTSSLEY 12
```

THIS PAGE BLANK (USPTO)

 WISEMAN (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSPRCH_PP protein - protein database search, using Smith-Waterman algorithm

on: Thu Sep 2 12:47:55 1999; Maspar time 4.49 Seconds
 Modular output not generated. 145.743 Million cell updates/sec

Title: >US-08-599-226-33
 Description: (1-12) from US08599226.pep
 Perfect Score: 74
 Sequence: 1 ASFLSTSSLEY 12

Scoring table: PAM 150
 Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: splemb19
 1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
 5:sp-invertebrate 6:sp-mammal 7:sp-mic 8:sp-organelle
 9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
 13:sp-invertebrate 14:sp-virus

Statistics: Mean 24.159; Variance 28.834; scale 0.838

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	55	74.3	475	13	093514	AXIAL PROTOCADHERIN (F
2	54	73.0	469	2	032204	YVSH PROTEIN.
3	54	73.0	1035	13	057537	NE-PROTOCADHERIN.
4	54	73.0	1069	4	060245	PCDH7 (BH-PCDH).A.
5	54	73.0	1069	11	088185	BH-PROTOCADHERIN-A.
6	54	73.0	1072	4	060246	PCDH7 (BH-PCDH).B.
7	54	73.0	1200	4	060247	PCDH7 (BH-PCDH).C.
8	53	71.6	112	3	007255	INTERNAL ORF OF L3149
9	52	70.3	809	5	020702	SIMILAR TO MATRIN F/G.
10	52	70.3	939	4	015071	KIAA0365 (FRAGMENT).
11	52	70.3	949	4	060369	KIAA0365 (FRAGMENT).
12	52	70.3	1026	4	008174	PROTOCADHERIN 42 PRECU
13	51	68.9	649	3	060167	PROTEIN COMPLEX ASSEMB
14	50	67.6	60	14	069367	GLYCOPROTEIN GG (FRAGM
15	50	67.6	198	4	075805	HOKA-9A.
16	50	67.6	296	5	026364	HKB-SEGMENTATION GAP-G
17	50	67.6	311	13	099162	MELANOMA RECEPTOR PROT
18	50	67.6	346	13	P11776	PROTO-ONCOGENE RECEPTO
19	50	67.6	471	5	026199	SERA-3 (FRAGMENT).
20	50	67.6	503	2	026074	PROTEIN-EXPORT MEMBRAN

21	50	67.6	509	11	061009	SCAVENGER RECEPTOR CIA	1.34e+01
22	50	67.6	509	11	P97943	SCAVENGER RECEPTOR CIA	1.34e+01
23	50	67.6	509	11	018824	TYPE II PNEUMOCYTE CD4	1.34e+01
24	50	67.6	509	11	088548	PHENOLHYDROXYLASE COMP	1.34e+01
25	50	67.6	511	2	043881	DMS OXYGENASE COMPONENT	1.34e+01
26	50	67.6	511	2	032431	GLYCOPROTEIN G (HOMOLO	1.34e+01
27	50	67.6	593	14	087093	REPLICATION PROTEIN E1	1.34e+01
28	50	67.6	599	14	080929	BAC IG005110.	1.34e+01
29	50	67.6	662	10	023064	HYPOPHOSPHATE 18.1 KD P	2.13e+01
30	49	66.2	155	2	051505	SIMILAR TO PRCV-1 SERI	2.13e+01
31	49	66.2	303	14	084594	UVRA PROTEIN.	2.13e+01
32	49	66.2	442	2	050151	CD36-RELATED CLASS B S	2.13e+01
33	49	66.2	509	11	060417	PHENOL HYDROXYLASE P3	2.13e+01
34	49	66.2	516	2	052173	FATTY ACID TRANSPORTER	2.13e+01
35	49	66.2	643	3	042633	F8A5.30 PROTEIN.	2.13e+01
36	49	66.2	664	10	022716	SIMILAR TO LIN-15B PRO	2.13e+01
37	49	66.2	890	5	020388	PUTATIVE EXTRACELLULAR	3.37e+01
38	49	66.2	1476	13	090285	HSPI8 TRANSCRIPTION RE	3.37e+01
39	48	64.9	224	2	008247	MO182.5 PROTEIN.	3.37e+01
40	48	64.9	349	5	017959	HYPOPHOSPHATE 42.6 KD P	3.37e+01
41	48	64.9	376	2	085449	(SUBCLONE PAWIT218) AM	3.37e+01
42	48	64.9	500	10	041537	INTEGRIN BETA3.	3.37e+01
43	48	64.9	781	13	092071	SIMILAR TO LIGAND-GATE	3.37e+01
44	48	64.9	932	5	001623	CHROMOSOME XVI READING	3.37e+01
45	48	64.9	1495	3	012280		

ALIGNMENTS

RESULT ID	Score	Query Match	Length	DB ID	Description	Pred. No.
1	55	74.3	475	13	093514	AXIAL PROTOCADHERIN (F
2	54	73.0	469	2	032204	YVSH PROTEIN.
3	54	73.0	1035	13	057537	NE-PROTOCADHERIN.
4	54	73.0	1069	4	060245	PCDH7 (BH-PCDH).A.
5	54	73.0	1069	11	088185	BH-PROTOCADHERIN-A.
6	54	73.0	1072	4	060246	PCDH7 (BH-PCDH).B.
7	54	73.0	1200	4	060247	PCDH7 (BH-PCDH).C.
8	53	71.6	112	3	007255	INTERNAL ORF OF L3149
9	52	70.3	809	5	020702	SIMILAR TO MATRIN F/G.
10	52	70.3	939	4	015071	KIAA0365 (FRAGMENT).
11	52	70.3	949	4	060369	KIAA0365 (FRAGMENT).
12	52	70.3	1026	4	008174	PROTOCADHERIN 42 PRECU
13	51	68.9	649	3	060167	PROTEIN COMPLEX ASSEMB
14	50	67.6	60	14	069367	GLYCOPROTEIN GG (FRAGM
15	50	67.6	198	4	075805	HOKA-9A.
16	50	67.6	296	5	026364	HKB-SEGMENTATION GAP-G
17	50	67.6	311	13	099162	MELANOMA RECEPTOR PROT
18	50	67.6	346	13	P11776	PROTO-ONCOGENE RECEPTO
19	50	67.6	471	5	026199	SERA-3 (FRAGMENT).
20	50	67.6	503	2	026074	PROTEIN-EXPORT MEMBRAN

Result	Score	Query Match	Length	DB ID	Description	Pred. No.
1	55	74.3	475	13	093514	AXIAL PROTOCADHERIN (F
2	54	73.0	469	2	032204	YVSH PROTEIN.
3	54	73.0	1035	13	057537	NE-PROTOCADHERIN.
4	54	73.0	1069	4	060245	PCDH7 (BH-PCDH).A.
5	54	73.0	1069	11	088185	BH-PROTOCADHERIN-A.
6	54	73.0	1072	4	060246	PCDH7 (BH-PCDH).B.
7	54	73.0	1200	4	060247	PCDH7 (BH-PCDH).C.
8	53	71.6	112	3	007255	INTERNAL ORF OF L3149
9	52	70.3	809	5	020702	SIMILAR TO MATRIN F/G.
10	52	70.3	939	4	015071	KIAA0365 (FRAGMENT).
11	52	70.3	949	4	060369	KIAA0365 (FRAGMENT).
12	52	70.3	1026	4	008174	PROTOCADHERIN 42 PRECU
13	51	68.9	649	3	060167	PROTEIN COMPLEX ASSEMB
14	50	67.6	60	14	069367	GLYCOPROTEIN GG (FRAGM
15	50	67.6	198	4	075805	HOKA-9A.
16	50	67.6	296	5	026364	HKB-SEGMENTATION GAP-G
17	50	67.6	311	13	099162	MELANOMA RECEPTOR PROT
18	50	67.6	346	13	P11776	PROTO-ONCOGENE RECEPTO
19	50	67.6	471	5	026199	SERA-3 (FRAGMENT).
20	50	67.6	503	2	026074	PROTEIN-EXPORT MEMBRAN

RA MEDLINE: 98044033.
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
 RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
 RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,
 RA BROUILLET S., BRUGHI C.V., CALDWELL B., CAPANO V., CARTER N.M.,
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
 RA DENICOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMERSON P.T.,
 RA ENLIN K.D., EHRINGTON J., FABRET C., FERRARI E., FOULDER D.,
 RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,
 RA GHIM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,
 RA GIUSEPPI G., GUY B.J., HAGA K., HALECH J., HARWOOD C.R., HENAUT A.,
 RA HILBERT H., HOLZAPPEL S., HOSONO S., HULLO M.F., ITOYA M., JONES L.,
 RA JORIS B., KARAMATA D., KASAHARA Y., KLAER-BLANCHARD M., KLEIN C.,
 RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,
 RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,
 RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,
 RA NOONE D., O'REILLY M., OGAMA K., OGIMARA A., ODEGA B., PARK S.H.,
 RA PARRO V., POHL T.M., PORTELELLA D., PORWOLIK S., PRESCOTT A.M.,
 RA PRESCAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,
 RA RIEGER M., RIVOLTA C., ROCHE B., ROCHE M., SADAIE Y.,
 SAITO T., SCANLAIN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,
 RA SEKICUCHI J., SEKOSKA A., SERO S.J., SERRO P., SHIN B.S., SOLDO B.,
 RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,
 RA TAKEUCHI M., TAMAKOSHI A., TANAKA T., TERRESTA P., TOGONDI A.,
 RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,
 RA VIARI A., WAMBUIT R., WEDLER E., WEDLER H., WEITZINGER T.,
 RA WINERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,
 RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;
 RT "The complete genome sequence of the gram-positive bacterium Bacillus
 subtilis".
 RT NATURE 390:249-256(1997).
 RL NATURE 390:249-256(1997).
 RN [12]
 RP SEQUENCE FROM N.A.

RC STRAIN-168;
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA WIPAT A., BRIGNELL C.S., GUY J.B., ROSE M., EMERSON P.T.,
 RA HARWOOD C.R.;
 RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: 299121; E1186022; -;
 DR EMBL: A223978; E1249790; -;
 DR PFAM: PF00324; aa-permeases; 1.
 SQ SEQUENCE 469 AA; 50258 MW; 49186162 CRC32;

Query Match 73.0%; Score 54; DB 2; Length 469;
 Best Local Similarity 45.3%; Pred. No. 1.96e+00;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 376 FLHTSTPLDY 386
 QY 2 FLSTSSLEY 12

RESULT 3
 ID 057537 PRELIMINARY; PRT: 1035 AA.
 AC 057537;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE NF-PROTOCOLADHERIN.
 GN NEPC.
 OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
 OC MESOBATACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BRADLEY R.S., ESPESETH A., KINTNER C.;
 RL CURR. BIOL. 0:0-0(1998).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).

DR EMBL: AF043643; G2852363; -;
 DR PROSITE: PS00232; CADHERIN; 6.
 KW CELL ADHESION; GLYCOPROTEIN; TRANSMEMBRANE; CALCIUM-BINDING; REPEAT.
 SQ SEQUENCE 1035 AA; 113713 MW; 7E4D3C4E CRC32;

Query Match 73.0%; Score 54; DB 13; Length 1035;
 Best Local Similarity 60.0%; Pred. No. 1.96e+00;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 458 FLHTSAPLDY 467
 QY 3 FLSTSSLEY 12

RESULT 4
 ID 060245 PRELIMINARY; PRT: 1069 AA.
 AC 060245;
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE PCDH7 (BH-PCDH A).
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
 OC CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA YOSHIDA K., YOSHITOMO-NAKAGAWA K., SEKI N., SASAKI M., SUGANO S.;
 RL GENOMICS 0:0-0(1998).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 DR EMBL: AB006755; D1026122; -;
 DR PROSITE: PS00232; CADHERIN; 6.
 KW CELL ADHESION; GLYCOPROTEIN; TRANSMEMBRANE; CALCIUM-BINDING; REPEAT.
 SQ SEQUENCE 1069 AA; 116104 MW; F1732B30 CRC32;

Query Match 73.0%; Score 54; DB 4; Length 1069;
 Best Local Similarity 60.0%; Pred. No. 1.96e+00;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 486 FLHTSTPLDY 495
 QY 3 FLSTSSLEY 12

RESULT 5
 ID 088185 PRELIMINARY; PRT: 1069 AA.
 AC 088185;
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE BH-PROTOCOLADHERIN-A.
 GN PCDH7.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
 OC SCIROGNATHI; MORIDAE; MORINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 98277460.
 RA YOSHIDA K., YOSHITOMO-NAKAGAWA K., SEKI N., SASAKI M., SUGANO S.;
 RT "Cloning, expression analysis, and chromosomal localization of
 BH-protocadherin (PCDH7), a novel member of the cadherin
 superfamily".
 RT GENOMICS 49:458-461(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA YOSHIDA K., HIDA M., WATANABE M., YAMAGUCHI R., TATEYAMA S.,
 RA SUGANO S.;
 RT "cDNA cloning and chromosomal mapping of mouse BH-protocadherin.";
 RT DNA SEQ. 0:0-0(1998).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 DR EMBL: AB006758; D1033562; -;
 DR PROSITE: PS00232; CADHERIN; 5.
 KW CELL ADHESION; GLYCOPROTEIN; TRANSMEMBRANE; CALCIUM-BINDING; REPEAT.
 SQ SEQUENCE 1069 AA; 116314 MW; 0F3F60C6 CRC32;

Db 297 GFLGASVSLOY 307
 OY 2 SFLSTSSSLEY 12

RESULT 10
 ID 015071 PRELIMINARY: PRT: 939 AA.

AC 015071.
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE KIAA0365 (FRAGMENT).

OS HOMO SAPIENS (HUMAN).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
 CC CATARRHINI; HOMINIDAE; HOMO.

[1]
 RN SEQUENCE FROM N.A.

RC TISSUE-BRAIN;
 RX MEDLINE: 97349984.

NAKASE T., ISHIKAWA K., NAKAJIMA D., OHIRA M., SEKI N., MIYAJIMA N.,
 TANAKA A., KOTANI H., NOMURA N., OHARA O.;

"Prediction of the coding sequences of unidentified human genes. VII.
 The complete sequences of 100 new cDNA clones from brain which can
 code for large proteins in vitro."

RL DNA RES. 4:141-150(1997).
 DR EMBL: AB002363; D1021661; -.

FT NON-TER 1
 SO SEQUENCE 939 AA; 103677 MW; 4B75FDB9 CRC32;

Query Match 70.3%; Score 52; DB 4; Length 939;
 Best Local Similarity 70.0%; Pred. No. 5.19e+00;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 475 FLSDENSLEY 484
 OY 3 FLSTSSSLEY 12

RESULT 11
 ID 060369 PRELIMINARY: PRT: 949 AA.

AC 060369.
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE KIAA0365 (FRAGMENT).

OS HOMO SAPIENS (HUMAN).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
 CC CATARRHINI; HOMINIDAE; HOMO.

[1]
 RN SEQUENCE FROM N.A.

LAHERDIN J.E., MCCREARY P.M., SKORONSKI E., ADAMSON A.W.,
 BURKHART-SCHULTZ K., GORDON L., KYLE A., RAMIREZ M., STIMMSEN S.,

PHAN H., VELASCO N., GARNES J., DANGANAN L., POUNDSTONE P.,
 CHRISTENSEN M., GEORGESCU A., AVILA J., LIU S., ATTIX C., ANDREISE T.,

TRANKHEIM M., AMICO-KELLER G., COFIELD J., DUARTE S., LUCAS S.,
 BRICE R., THOMAS P., QUAN G., KRONMILLER B., ARELLANO A.,

MONTGOMERY M., OW D., NOLAN M., TRONG S., KOBAYASHI A., OLSEN A.O.,
 CARANO A.V.;

RA SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 RL EMBL: AC004447; G2978447; -.

FT NON-TER 1
 SO SEQUENCE 949 AA; 105111 MW; 1C0E3FB3 CRC32;

Query Match 70.3%; Score 52; DB 4; Length 949;
 Best Local Similarity 70.0%; Pred. No. 5.19e+00;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 475 FLSDENSLEY 484
 OY 3 FLSTSSSLEY 12

RESULT 12
 ID 008174 PRELIMINARY: PRT: 1026 AA.

AC 008174.
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE PROTOCADHERIN 42 PRECURSOR (PC42) (CADHERIN-LIKE PROTEIN).

OS HOMO SAPIENS (HUMAN).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
 CC CATARRHINI; HOMINIDAE; HOMO.

[1]
 RN SEQUENCE FROM N.A.

RC TISSUE-BRAIN;
 RX MEDLINE: 93285094.

SAO K., TANIHARA H., HEIMARK R.L., OBATA S., DAVIDSON M., ST JOHN T.,
 TAKEKUNI S., SUZUKI S.;

"Protocadherins: a large family of cadherin-related molecules in
 central nervous system."

EMBO J. 12:2249-2256(1993).
 RL -1- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTION PROCESSES AND
 IN CELL ADHESION.

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC FOUND AT CELL-CELL BOUNDARIES AND PROBABLY AT CELL-MATRIX
 CC BOUNDARIES.

CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE BRAIN AND NEURO-GLIAL
 CC CELLS.

CC -1- DEVELOPMENTAL STAGE: HIGHEST EXPRESSION IN ADULTS.

CC -1- SIMILARITY: BELONGS TO THE CADHERIN FAMILY OF CELL ADHESION
 CC MOLECULES. CONTAINS 7 CADHERIN-TYPE REPEATS.

CC -1- ALTERNATIVE PRODUCTS: DIFFERENT FORMS OF THIS PROTEIN THAT DIFFER
 CC AT THE C-TERMINUS ARE PRODUCED BY ALTERNATIVE SPLICING.

DR EMBL: L11370; G387675; -.

DR EMBL: L11369; G387673; -.

PFAM: PF00028; cadherin: 7.
 KW CELL ADHESION; TRANSMEMBRANE; CALCIUM-BINDING; ALTERNATIVE SPLICING;
 KW REPEAT; SIGNAL.

FT SIGNAL 1
 FT CHAIN 1
 FT DOMAIN 2
 FT TRANSMEM 819
 FT DOMAIN 840

FT REPEAT 42
 FT REPEAT 147
 FT REPEAT 147
 FT REPEAT 247
 FT REPEAT 354
 FT REPEAT 472
 FT REPEAT 473
 FT REPEAT 579
 FT REPEAT 681
 FT REPEAT 818
 FT VARSPLIC 1000

FT POTENTIAL. 2
 FT PROTOCADHERIN 42.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).

FT CADHERIN 1.
 FT CADHERIN 2.
 FT CADHERIN 3.
 FT CADHERIN 4.
 FT CADHERIN 5.
 FT CADHERIN 6.
 FT CADHERIN 7.
 FT VGFOLSTROPPLPHYGAIWTEWE ->
 LPHRYTFSA TSQAQLDPSQHSYDSGLE
 (IN PC42-8).

FT SEQUENCE 1026 AA; 111270 MW; 84A6E132 CRC32;

Query Match 70.3%; Score 52; DB 4; Length 1026;
 Best Local Similarity 50.0%; Pred. No. 5.19e+00;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 423 FL0TTPIDY 432
 OY 3 FLSTSSSLEY 12

RESULT 13
 ID 060167 PRELIMINARY: PRT: 649 AA.

AC 060167.
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE PROTEIN COMPLEX ASSEMBLY PROTEIN.

OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
 CC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASTOMYCETES;
 CC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE.

OC SCHIZOSACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA BECK A., REINHARDT R., WOOD V., RAJANDREAM M.A., BARRELL B.G.;
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AL023594; E1293401; -;
SQ SEQUENCE 649 AA; 72985 MW; 9C207DB2 CRC32;

Query Match 68.9%; Score 51; DB 3; Length 649;
Best Local Similarity 54.5%; Pred. No. 8.37e+00;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 234 NY1STARSLF 244
:||||:|
QY 2 SFLSTSSLEY 12

RESULT 14
069367; PRELIMINARY; PRT; 60 AA.
069367;
01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE GLYCOPROTEIN GG (FRAGMENT).
OS SIMIAN HERPESVIRUS S48.
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
OC ALPHAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93298054.
RA EBERLE R., ZHANG M., BLACK D.;
RT "Gene mapping and sequence analysis of the unique short region of the
RT simian herpesvirus SA 8 genome."
RL ARCH. VIROL. 130:391-411(1993).
DR EMBL: L05608; G331077; -;
FT NON_TER 1
SQ SEQUENCE 60 AA; 6623 MW; 5CE36ED8 CRC32;

Query Match 67.6%; Score 50; DB 14; Length 60;
Best Local Similarity 36.4%; Pred. No. 1.34e+01;
Matches 4; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Db 4 FELTASPGDLF 14
:|:|:|:|:|
QY 2 SFLSTSSLEY 12

SULT 15
AD 075805; PRELIMINARY; PRT; 198 AA.
075805;
01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HOXA-9A.
GN HOXA-9.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98290548.
RA KIM M.H., CHANG H.H., SHIN C., CHO M., PARK D., PARK H.W.;
RT "Genomic structure and sequence analysis of human HOXA-9.";
RL DNA CELL BIOL. 17:407-414(1998).
-|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: U81511; G3237298; -;
DR PROSITE: PS00027; HOMEBOX 1; 1.
KM HOMEBOX: NUCLEAR PROTEIN; DNA-BINDING.
SQ SEQUENCE 198 AA; 22874 MW; 307982CA CRC32;

Query Match 67.6%; Score 50; DB 4; Length 198;
Best Local Similarity 41.7%; Pred. No. 1.34e+01;

Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
Db 42 SFLSTGTLF 53
:||||:|
QY 1 ASFLSTSSLEY 12

Search completed: Thu Sep 2 12:48:16 1999
Job time : 21 secs.

THIS PAGE BLANK (USPTO)

OC ALPHAHERPESVIRINAE: SIMPLEXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 85160822.
 RA MCGEECH D.J., DOLAN A., DONALD S., RIXON F.J.:
 RT "Sequence determination and genetic content of the short unique
 RL region in the genome of herpes simplex virus type 1.";
 J. MOL. BIOL. 181:1-13(1985).
 CC -1- THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND 2: GH, GB, GC,
 CC GG, GD, GI, AND GE.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L00036; G291500; -
 CC EMBL: X14112; G59562; -
 CC EMBL: X14112; E312376; -
 CC EMBL: X02138; G59870; -
 CC DR PIR: A05239; Q0BE74.
 CC DR GLYCOPROTEIN.
 CC KW CARBOHYD 28 28 POTENTIAL.
 CC FT CARBOHYD 49 49 POTENTIAL.
 CC FT CARBOHYD 49 49 POTENTIAL.
 CC SQ SEQUENCE 238 AA; 25238 MW; 284A50FD CRC32:
 CC
 CC Query Match 63.5%; Score 47; DB 1; Length 238;
 CC Best Local Similarity 50.0%; Pred. No. 2.36e+01;
 CC Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Db 180 SFLSTASPLD 189
 CC QY 111:11:11:
 CC 2 SFLSTSSLE 11
 CC
 CC RESULT 15
 CC ID PH4H.PSEAE STANDARD; PRT; 262 AA.
 CC AC P43334;
 CC DT 01-NOV-1995 (REL. 32, CREATED)
 CC DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 CC DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 CC DE PHENYLALANINE-4-HYDROXYLASE (EC 1.14.16.1) (PAH) (PHE-4-
 CC MONOOXYGENASE).
 CC GN PHHA.
 CC PSEUDOMONAS AERUGINOSA.
 CC BACTERIA: PROTEOBACTERIA: GAMMA SUBDIVISION; PSEUDOMONAS GROUP;
 CC PSEUDOMONAS.
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN-ATCC 15692 / PA01;
 CC RX MEDLINE: 94151331.
 CC RA ZHAO G., XIA T., SONG J., ROY R.A.:
 CC RT "Pseudomonas aeruginosa possesses homologues of mammalian
 CC phenylalanine hydroxylase and 4 alpha-carbinolamine dehydratase/DCOH
 CC as part of a three-component gene cluster.";
 CC PROC. NATL. ACAD. SCI. U.S.A. 91:1366-1370(1994).
 CC RL
 CC -1- CATALYTIC ACTIVITY: L-PHENYLALANINE + TETRAHYDROBIOPTERIN + O(2) =
 CC L-TYROSINE + DIHYDROBIOPTERIN + H(2)O.
 CC -1- COFACTOR: THIS ENZYME REQUIRES A FERROUS ION
 CC -1- PATHWAY: RATE-LIMITING STEP IN PHENYLALANINE CATABOLISM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO THE BIOPTERIN-DEPENDENT AROMATIC AMINO ACID
 CC HYDROXYLASES FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M88627; G476741; -
 CC DR PROSITE; PS00367; BIOPTERIN-HYDROXYL; 1.
 CC DR PFAM; PF00351; bioplerin_H; 1.
 CC DR HSSP; P04177; 1TCH.
 CC KW OXIDOREDUCTASE; MONOOXYGENASE; PHENYLALANINE CATABOLISM; IRON.
 CC FT METAL 121 121 IRON (BY SIMILARITY).
 CC FT METAL 126 126 IRON (BY SIMILARITY).
 CC FT METAL 126 126 IRON (BY SIMILARITY).
 CC SQ SEQUENCE 262 AA; 30288 MW; 6D464AD4 CRC32;
 CC
 CC Query Match 63.5%; Score 47; DB 1; Length 262;
 CC Best Local Similarity 41.7%; Pred. No. 2.36e+01;
 CC Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC Db 102 ATRFRPELDY 113
 CC QY 111:11:11:
 CC 1 ASFLSTSSLEY 12

Search completed: Thu Sep 2 12:47:38 1999
 Job time : 8 secs.

This Page Blank (uspto)

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U50987; G1575011. -
KM TRICARBOXYLIC ACID CYCLE: ELECTRON TRANSPORT; HEME; TRANSMEMBRANE;
KM MITOCHONDRION; TRANSIT PEPTIDE.

FT CHAIN 1 55 MITOCHONDRION (POTENTIAL).
FT TRANSIT 56 158 CYTOCHROME B SMALL SUBUNIT.
FT TRANSMEM 70 90 POTENTIAL.
FT TRANSMEM 125 141 POTENTIAL.
SO SEQUENCE 158 AA; 17096 MW; 703D5238 CRC32;

Query Match
Best Local Similarity 33.3%; Pred. No. 1.47e+01;
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

81 AAYLPCSAMDY 92
1 ASFLSTSSSLEY 12

RESULT 12
ID DHSD_HUMAN STANDARD; PRT; 159 AA.

DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE SUCCINATE DEHYDROGENASE [UBIQUINONE] CYTOCHROME B SMALL SUBUNIT
DE PRECURSOR (CYBS) (SUCCINATE-UBIQUINONE REDUCTASE MEMBRANE ANCHOR
DE SUBUNIT)
CN SDSD OR SDH4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.

[1]
RN SEQUENCE FROM N.A.
RC TISSUE-LIVER.
RA MEDLINE; 98194224.
RX HIRAWAKE H., TANIMAKI M., KIJIMA S., KIRA K.;
RT "Cytochrome b in human complex II (succinate-ubiquinone
oxidoreductase): cDNA cloning of the components in liver mitochondria
and chromosome assignment of the genes for the large (SDHC) and small
(SDHB) subunits to 1q21 and 11q23";
RL CYTOGENET. CELL GENET. 79:132-138(1997).

-1 SUBUNIT: COMPOSED OF A 27 KD IRON PROTEIN (IP), A 70 KD
FLAVOPROTEIN (FP) AND A CYTOCHROME B COMPOSED OF TWO INTEGRAL
MEMBRANE PROTEINS.
-1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
INNER MEMBRANE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AB006202; D1022913; -
DR MIM; 602690; -

KM TRICARBOXYLIC ACID CYCLE: ELECTRON TRANSPORT; HEME; TRANSMEMBRANE;
KM MITOCHONDRION; TRANSIT PEPTIDE.

FT CHAIN 1 56 MITOCHONDRION (POTENTIAL).
FT TRANSIT 57 159 CYTOCHROME B SMALL SUBUNIT.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 126 142 POTENTIAL.
SO SEQUENCE 159 AA; 17043 MW; F4221825 CRC32;

Query Match
64.9%; Score 48; DB 1; Length 159;

Best Local Similarity 33.3%; Pred. No. 1.47e+01;
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Db 82 AAYLPCSAMDY 93
1 ASFLSTSSSLEY 12

RESULT 13
ID YOE1_STRAF STANDARD; PRT; 229 AA.

DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN IN OLEO 5' REGION (ORF1) (FRAGMENT).
OS STREPTOMYCES ANTIBIOTICUS.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIAE;
OC ACTINOMYCETALES; STREPTOMYCINAE; STREPTOMYCETACEAE; STREPTOMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 11891.
RA MEDLINE; 94063510.
RX HERNANDEZ C., OLANO C., MENDEZ C., SALAS J.A.;
RT "Characterization of a Streptomyces antibiotic gene cluster
encoding a glycosyltransferase involved in oleandomycin
inactivation";
RL GENE 134:139-140(1993).

-1 FUNCTION: MAY PARTICIPATE IN OLEANDOMYCIN SECRETION DURING
CC ANTI-IONIC PRODUCTION.
CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1 SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE MALFG
CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; 222577; G296489; -
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBR. PARTIAL.
DR PRM; PF00528; BPD_transp. 1.

KM HYPOTHETICAL PROTEIN; TRANSMEMBRANE; TRANSPORT.

FT NON_TER 1 1
FT TRANSMEM 26 42 POTENTIAL.
FT TRANSMEM 63 79 POTENTIAL.
FT TRANSMEM 94 110 POTENTIAL.
FT TRANSMEM 141 157 POTENTIAL.
FT TRANSMEM 198 214 POTENTIAL.
SO SEQUENCE 229 AA; 24312 MW; EF5F0F21 CRC32;

Query Match
Best Local Similarity 50.0%; Pred. No. 2.35e+01;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 199 AAYLPCSAMDY 210
1 ASFLSTSSSLEY 12

RESULT 14
ID VGLG_HSV11 STANDARD; PRT; 238 AA.

AC P06484;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE GLYCOPROTEIN G.
GG OR US4.
OS HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN 17).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;

Best Local Similarity 60.0%; Pred. No. 9.09e+00;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 8 FLEVSFLEY 17
1 : 1:111
QY 3 FLETSLSLEY 12

RESULT 9
ID CYP2A97 STANDARD; PRT; 439 AA.
AC P22A97;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CYSTEINE PROTEINASE PRECURSOR (EC 3.4.22.-).
OS THEILERIA PARVA.
OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; PIROPLASMDA; THEILERIIDAE;
OC THEILERIA.
RN [1]
SEQUENCE FROM N.A.
MEDLINE: 91009278.
NENE V., GOBRIGHT E., MUSOKE A.J., LONSDALE-ECCLES J.D.;
"A single exon codes for the enzyme domain of a protozoan cysteine
protease."
RT J. BIOL. CHEM. 265:18047-18050(1990).
RN [2]
SEQUENCE OF 178-439 FROM N.A.
RC STRAIN-MUGUGA;
RX MEDLINE: 92228011.
RA NENE V., IAMS K.P., GOBRIGHT E., MUSOKE A.J.;
RT "Characterisation of the gene encoding a candidate vaccine antigen of
Theileria parva sporozoites."
RL MOL. BIOCHEM. PARASITOL. 51:17-27(1992).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
PAPAIN FAMILY OF THIOL PROTEASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M37791; G161873; -
DR EMBL: M67476; G161879; -
DR PIR: A36083; KHOBT.
PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
PROSITE: PS00639; THIOL_PROTEASE_HIS; 1.
PROSITE: PS00640; THIOL_PROTEASE_ASN; 1.
PFAM: PF00112; Cys-protease; 1.
DR HSP: P14080; 1VAL.
KW HYDROLASE; THIOL PROTEASE; ZYMOGEN; SIGNAL; GLYCOPROTEIN.
FT SIGNAL 1 59
FT PROPEP 60 228
FT CHAIN 229 439
FT ACT_SITE 252 252
FT ACT_SITE 381 381
FT ACT_SITE 403 403
FT DISULFID 249 290
FT DOMAIN 165 181
FT CARBOHYD 205 205
FT SEQUENCE 439 AA; 50179 MW; B4BD9301 CRC32;
POTENTIAL.
ENZYMES (BY SIMILARITY).

Query Match 66.2%; Score 49; DB 1; Length 439;
Best Local Similarity 54.5%; Pred. No. 9.09e+00;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 108 GFLSDPKLEY 118
1 : 111 : 111
QY 2 SFLSTSSLEY 12

RESULT 10
ID Y9A_SCHPO STANDARD; PRT; 530 AA.
AC 009788;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 54.2 KD SERINE-RICH PROTEIN C1366.10C IN CHROMOSOME I
DE PRECURSOR.
GN SPAC1366.10C.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
RN [1]
SEQUENCE FROM N.A.
RC STRAIN-9712;
RA ODELL C., BOWMAN S., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: 254308; G1008995; -
KW HYPOTHETICAL PROTEIN; SIGNAL.
FT SIGNAL 1 18
FT CHAIN 19 530
FT CARBOHYD 55 55
FT CARBOHYD 120 120
FT CARBOHYD 128 128
FT SEQUENCE 530 AA; 54210 MW; 48D83906 CRC32;
POTENTIAL.

Query Match 66.2%; Score 49; DB 1; Length 530;
Best Local Similarity 54.5%; Pred. No. 9.09e+00;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 11 AFLSTVGALPY 21
1 : 111 : 111
QY 2 SFLSTSSLEY 12

RESULT 11
ID Q95123 STANDARD; PRT; 158 AA.
AC Q95123;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE SUCCINATE DEHYDROGENASE (UBIQUINONE) CYTOCHROME B SMALL SUBUNIT
DE PRECURSOR (CYBS) (SUCCINATE-UBIQUINONE REDUCTASE MEMBRANE ANCHOR
DE SUBUNIT) (OPS3).
DE SDHD OR SDH4.
GN BOS TAURUS (BOVINE).
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINA; BOS.
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA SHENOY S.K.; YU L.; YU C.A.;
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SUBUNIT: COMPOSED OF A 27 KD IRON PROTEIN (1P), A 70 KD
CC FLAVOPROTEIN (FP) AND A CYTOCHROME B COMPOSED OF TWO INTEGRAL
CC MEMBRANE PROTEINS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

FT REPEAT 387 422 1.
 FT REPEAT 423 471 2.
 FT REPEAT 472 519 3.
 FT REPEAT 520 566 4.
 SO SEQUENCE 579 AA: 60543E32 CRC32:

Query Match 67.6%; Score 50; DB 1; Length 579;
 Best Local Similarity 50.0%; Pred. No. 5.57e+00;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 133 AAFLOWSTALDY 144
 1:11 1:11
 Qy 1 AAFLOWSTALDY 12

RESULT 7
 ID XMRK_XIPMA STANDARD: PRT: 1166 AA.
 AC P13388;
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 37, LAST ANNOTATION UPDATE)
 G1 MELANOMA RECEPTOR PROTEIN-TYROSINE KINASE PRECURSOR (EC 2.7.1.112).
 G1 XMRK OR 10.
 OS XIPHOPHORUS MACULATUS (SOUTHERN PLATYFISH).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
 OC TELEOSTEI; EUTELEOSTEI; ACANTHOPTERYGII; ATHERINOMORPHA;
 OC CYRINODONTIFORMES; CYRINODONTIIDEI; POECILIIDAE; XIPHOPHORUS.
 RN [1]
 RX SEQUENCE FROM N.A.
 RX MEDLINE: 90015140.
 RA WITTBRODT J., ADAM D., MALITSCHER B., MAUELER W., RAULF F.,
 RA TELLING A., ROBERTSON S.M., SCHARL M.;
 RT "Novel putative receptor tyrosine kinase encoded by the melanoma-
 inducing Tu locus in xiphophorus.";
 RL NATURE 341:415-421(1989).
 CC -1 FUNCTION: PROBABLE RECEPTOR WITH PROTEIN-TYROSINE KINASE ACTIVITY.
 CC -1 CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +
 CC PROTEIN TYROSINE PHOSPHATE.
 CC -1 SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1 DISEASE: INVOLVED IN PIGMENT CELLS MALIGNANT MELANOMAS.
 CC -1 SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -1
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X16891; G65291; -
 DR PIR: S06142; S06142;
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP. 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR. 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM. 1.
 DR PFAM: PF00069; PKINASE. 1.
 DR PFAM: PF00757; Furin-like. 1.
 DR PFAM: PF01030; Recep_L-domain. 2.
 DR HSPR: P1162; 1ref.
 DR TRANSFERRIN; GLYCOPROTEIN; DUPLICATION; RECEPTOR; SIGNAL;
 KW TRANSFERRIN; TYROSINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION;
 KW PROTO-ONCOGENE.
 FT SIGNAL 1 25
 FT CHAIN 26 1166 MELANOMA RECEPTOR TYROSINE KINASE.
 FT DOMAIN 26 642 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 643 665 POTENTIAL.
 FT DOMAIN 666 1166 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 709 976 PROTEIN KINASE.
 FT NP_BIND 715 723 ATP (BY SIMILARITY).
 FT BINDING 742 742 ATP (BY SIMILARITY).
 FT ACT_SITE 834 834 BY SIMILARITY.
 FT CARBOHYD 114 114 POTENTIAL.
 FT CARBOHYD 144 144 POTENTIAL.

FT CARBOHYD 201 201 POTENTIAL.
 FT CARBOHYD 356 356 POTENTIAL.
 FT CARBOHYD 365 365 POTENTIAL.
 FT CARBOHYD 398 398 POTENTIAL.
 FT CARBOHYD 417 417 POTENTIAL.
 FT CARBOHYD 501 501 POTENTIAL.
 FT CARBOHYD 575 575 POTENTIAL.
 FT CARBOHYD 620 620 POTENTIAL.
 SO SEQUENCE 1166 AA: 129877 MW: 48F8E7C0 CRC32:

Query Match 67.6%; Score 50; DB 1; Length 1166;
 Best Local Similarity 50.0%; Pred. No. 5.57e+00;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 1145 FLPAENLEY 1154
 11:11:11
 Qy 3 FLPAENLEY 12

RESULT 8
 ID RIR2_MYCPN STANDARD: PRT: 339 AA.
 AC P75461;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE BETA CHAIN (EC 1.17.4.1)
 DE (RIBONUCLEOTIDE REDUCTASE).
 GN NRDE.
 OS MYCOPLASMA PNEUMONIAE.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLICUTES;
 OC MYCOPLASMATACEAE; MYCOPLASMA.
 RN [1]
 RX SEQUENCE FROM N.A.
 RX STRAIN-ATCC 29342 / M129;
 RX MEDLINE: 97105885.
 RA HIMMELREICH R., HILBERT H., PLAGENS H., PIKLE E., LI B.-C.,
 RA HERMANN R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae.";
 RL NUCLEIC ACIDS RES. 24:4420-4449(1996).
 CC -1 FUNCTION: CATALYZES THE BIOSYNTHESIS OF DEOXYRIBONUCLEOTIDES FROM
 CC THE CORRESPONDING RIBONUCLEOTIDES, PRECURSORS THAT ARE NECESSARY
 CC FOR DNA SYNTHESIS (BY SIMILARITY).
 CC -1 CATALYTIC ACTIVITY: 2'-DEOXYRIBONUCLEOSIDE DIPHOSPHATE + OXIDIZED
 CC THIOREDOXIN + H(2)O = RIBONUCLEOSIDE DIPHOSPHATE + REDUCED
 CC THIOREDOXIN.
 CC -1 COFACTOR: CONTAINS TWO IRON IONS (BY SIMILARITY).
 CC -1 PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
 CC -1 SUBUNIT: Tetramer of two alpha and two beta chains
 CC (BY SIMILARITY).
 CC -1 SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
 CC SMALL CHAIN FAMILY. MORE SIMILAR TO ENTEROBACTERIAL NRDE THAN TO
 CC NRDB. SEEMS TO LACK TWO OF THE IRON-BINDING RESIDUES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000050; G1674208; -
 DR PROSITE: PS00366; RIBONUCLEOTIDE REDUCTASE; FALSE-NEG.
 DR PFAM: PF00268; RIBONUCLEOTIDE REDUCTASE; IRON.
 KW OXIDOREDUCTASE; DNA REPLICATION; IRON.
 FT METAL 87 87 IRON 1 (BY SIMILARITY).
 FT METAL 121 121 IRON 1 (BY SIMILARITY).
 FT METAL 215 215 IRON 2 (BY SIMILARITY).
 FT ACT_SITE 125 125 BY SIMILARITY.
 SO SEQUENCE 339 AA: 39413 MW: B73F2761 CRC32:
 Query Match 66.2%; Score 49; DB 1; Length 339;

```

RESULT 4
ID UFO2_MAIZE STANDARD: PRT: 471 AA.
AC P16165;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DE FLAVONOL 3-O-GLUCOSYLTRANSFERASE (EC 2.4.1.91) (UDP-GLUCOSE FLAVONOID
  3-O-GLUCOSYLTRANSFERASE) (BRONZE-1) (B2-MC2 ALLELE).
GN B21 OR UGT71A1.
OS ZEA MAYS (MAIZE).
OC EUPHYLLOPHYTES; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
  OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES;
  OC POACEAE; ZEA.
RN [1]
RP SEQUENCE FROM N.A.
RA FURTER D., SCHIEFELBEIN J.W., JOHNSTON F., NELSON O.E. JR.;
RT "Sequence comparisons of 3 wild-type bronze-1 alleles from Zea mays."
  PLANT MOL. BIOL. 11:473-481(1988).
  -1- FUNCTION: IN THE PRESENCE OF OTHER NECESSARY COLOR FACTORS, THIS
    GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
    PIGMENTS.
  CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + A FLAVONOL = UDP + FLAVONOL
  CC -1- 3-O-D-GLUCOSIDE.
  CC -1- PATHWAY: ONE OF THE ESSENTIAL AND TERMINAL STEPS IN THE
  CC ANTHOCYANIN BIOSYNTHETIC PATHWAY.
  CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
  CC -----
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
  CC -----
  DR EMBL: X13501; G295854; -.
  DR PIR: S08325; S08325.
  DR MAIZEDB: 13885; -.
  DR PROSITE: PS00375; UDCPT: 1.
  DR PFAM: PF00201; UDCPT: 2.
  KW TRANSFERASE; GLYCOSYLTRANSFERASE.
  SQ SEQUENCE 471 AA; 48621 MW; 3158C5E0 CRC32;

Query Match 67.6%; Score 50; DB 1; Length 471;
Best Local Similarity 88.9%; Pred. No. 5.57e+00;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1
OY 47 SFLSTASL 55
  11111111
  2 SFLSTSSL 10

RESULT 5
ID UFO3_MAIZE STANDARD: PRT: 471 AA.
AC P16167;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DE FLAVONOL 3-O-GLUCOSYLTRANSFERASE (EC 2.4.1.91) (UDP-GLUCOSE FLAVONOID
  3-O-GLUCOSYLTRANSFERASE) (BRONZE-1) (B2-W22 ALLELE).
GN B21 OR UGT71A1.
OS ZEA MAYS (MAIZE).
OC EUPHYLLOPHYTES; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
  OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES;
  OC POACEAE; ZEA.
RN [1]
RP SEQUENCE FROM N.A.
RA FURTER D., SCHIEFELBEIN J.W., JOHNSTON F., NELSON O.E. JR.;
RT "Sequence comparisons of 3 wild-type bronze-1 alleles from Zea mays."
  PLANT MOL. BIOL. 11:473-481(1988).
  -1- FUNCTION: IN THE PRESENCE OF OTHER NECESSARY COLOR FACTORS, THIS
    GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
    PIGMENTS.
  CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + A FLAVONOL = UDP + FLAVONOL
  CC -1- 3-O-D-GLUCOSIDE.
  CC -1- PATHWAY: ONE OF THE ESSENTIAL AND TERMINAL STEPS IN THE
  CC ANTHOCYANIN BIOSYNTHETIC PATHWAY.
  CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
  CC -----
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
  CC -----
  DR EMBL: X13502; G22506; -.
  DR PIR: X07937; G22210; -.
  DR PIR: S01037; S01037.
  DR PIR: S08326; S08326.
  DR MAIZEDB: 13885; -.
  DR PROSITE: PS00375; UDCPT: 1.
  DR PFAM: PF00201; UDCPT: 2.
  KW TRANSFERASE; GLYCOSYLTRANSFERASE.
  SQ SEQUENCE 471 AA; 48673 MW; 4A3C6193 CRC32;

Query Match 67.6%; Score 50; DB 1; Length 471;
Best Local Similarity 88.9%; Pred. No. 5.57e+00;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 47 SFLSTASL 55
  11111111
  2 SFLSTSSL 10

OY 2 SFLSTSSL 10

RESULT 6
ID YR47_CAEEL STANDARD: PRT: 579 AA.
AC O09563;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL 66.0 KD PROTEIN F47D12.7 IN CHROMOSOME III.
GN F47D12.7.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
  OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
  RN [1]
  RP SEQUENCE FROM N.A.
  RC STRAIN-BRISTOL N2;
  RA TAICH A.;
  RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
  CC -1- SIMILARITY: STRONG. TO C.ELEGANS T16H12.6 AND SOME, TO MOUSE M1PP.
  CC -----
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
  CC -----
  DR EMBL: U22831; G722359; -.
  DR MORMREP: F47D12.7; CE01951.
  DR PFAM: PF00651; BTB: 1.
  DR PFAM: PF01344; Kelch: 6.
  KW HYPOTHETICAL PROTEIN; REPEAT.
  FT DOMAIN 387 566 4 X APPROXIMATE TANDEM REPEATS.

```

```

RX MEDLINE; 88284304.
RA RALSTON E.J., ENGLISH J.J., DOONER H.K.;
RT "Sequence of three bronze alleles of maize and correlation with the
  RT genetic fine structure."
  RL GENETICS 119:185-197(1988).
  -1- FUNCTION: IN THE PRESENCE OF OTHER NECESSARY COLOR FACTORS, THIS
    GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
    PIGMENTS.
  CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + A FLAVONOL = UDP + FLAVONOL
  CC -1- 3-O-D-GLUCOSIDE.
  CC -1- PATHWAY: ONE OF THE ESSENTIAL AND TERMINAL STEPS IN THE
  CC ANTHOCYANIN BIOSYNTHETIC PATHWAY.
  CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
  CC -----
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
  CC -----
  DR EMBL: X13502; G22506; -.
  DR EMBL: X07937; G22210; -.
  DR PIR: S01037; S01037.
  DR PIR: S08326; S08326.
  DR MAIZEDB: 13885; -.
  DR PROSITE: PS00375; UDCPT: 1.
  DR PFAM: PF00201; UDCPT: 2.
  KW TRANSFERASE; GLYCOSYLTRANSFERASE.
  SQ SEQUENCE 471 AA; 48673 MW; 4A3C6193 CRC32;

Query Match 67.6%; Score 50; DB 1; Length 471;
Best Local Similarity 88.9%; Pred. No. 5.57e+00;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 SFLSTSSL 10

Db 47 SFLSTASL 55
  11111111
  2 SFLSTSSL 10

OY 2 SFLSTSSL 10

RESULT 6
ID YR47_CAEEL STANDARD: PRT: 579 AA.
AC O09563;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL 66.0 KD PROTEIN F47D12.7 IN CHROMOSOME III.
GN F47D12.7.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
  OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
  RN [1]
  RP SEQUENCE FROM N.A.
  RC STRAIN-BRISTOL N2;
  RA TAICH A.;
  RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
  CC -1- SIMILARITY: STRONG. TO C.ELEGANS T16H12.6 AND SOME, TO MOUSE M1PP.
  CC -----
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
  CC -----
  DR EMBL: U22831; G722359; -.
  DR MORMREP: F47D12.7; CE01951.
  DR PFAM: PF00651; BTB: 1.
  DR PFAM: PF01344; Kelch: 6.
  KW HYPOTHETICAL PROTEIN; REPEAT.
  FT DOMAIN 387 566 4 X APPROXIMATE TANDEM REPEATS.

```

 W O R L D
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Search: protein - protein database search, using Smith-Waterman algorithm
 on: Thu Sep 2 12:47:30 1999; Masparr time 2.21 Seconds
 Output not generated. 153.600 Million cell updates/sec

Title: >US-08-599-226-33
 Description: (1-12) from US08599226.pep
 Perfect Score: 74
 Sequence: 1 ASFLSTSSSLEY 12

Scoring table: PAM 150
 Gap 15

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot37
 I:swissprot

Statistics: Mean 25.085; Variance 27.334; scale 0.918

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

ult	Score	Query	Match	Length	ID	Description	Pred. No.
1	52	70.3	517	1	DMPN_PESP	PHENOL HYDROXYLASE P3	2.05e+00
2	50	67.6	464	1	WCAM_ECOLI	COLANIC ACID BIOSYNTHET	5.57e+00
3	50	67.6	471	1	UFOI_MAIZE	FLAVONOL 3-O-GLUCOSYL	5.57e+00
4	50	67.6	471	1	UFOI_MAIZE	FLAVONOL 3-O-GLUCOSYL	5.57e+00
5	50	67.6	471	1	UFOI_MAIZE	FLAVONOL 3-O-GLUCOSYL	5.57e+00
6	50	67.6	471	1	UFOI_MAIZE	FLAVONOL 3-O-GLUCOSYL	5.57e+00
7	50	67.6	471	1	UFOI_MAIZE	FLAVONOL 3-O-GLUCOSYL	5.57e+00
8	49	66.2	439	1	CYSP_THERA	HYPOHETICAL 34.2 KD S	1.47e+01
9	49	66.2	439	1	CYSP_THERA	HYPOHETICAL 34.2 KD S	1.47e+01
10	49	66.2	439	1	CYSP_THERA	HYPOHETICAL 34.2 KD S	1.47e+01
11	48	64.9	158	1	DHSD_BOVIN	SUCCINATE DEHYDROGENAS	1.47e+01
12	48	64.9	158	1	DHSD_BOVIN	SUCCINATE DEHYDROGENAS	1.47e+01
13	47	63.5	238	1	VGIC_HSV1	GLYCOPROTEIN G, HYDROX	2.36e+01
14	47	63.5	238	1	VGIC_HSV1	GLYCOPROTEIN G, HYDROX	2.36e+01
15	47	63.5	238	1	VGIC_HSV1	GLYCOPROTEIN G, HYDROX	2.36e+01
16	47	63.5	238	1	VGIC_HSV1	GLYCOPROTEIN G, HYDROX	2.36e+01
17	47	63.5	238	1	VGIC_HSV1	GLYCOPROTEIN G, HYDROX	2.36e+01
18	47	63.5	238	1	VGIC_HSV1	GLYCOPROTEIN G, HYDROX	2.36e+01
19	47	63.5	238	1	VGIC_HSV1	GLYCOPROTEIN G, HYDROX	2.36e+01
20	47	63.5	238	1	VGIC_HSV1	GLYCOPROTEIN G, HYDROX	2.36e+01
21	47	63.5	238	1	VGIC_HSV1	GLYCOPROTEIN G, HYDROX	2.36e+01
22	46	62.2	207	1	CLPP_ECOLI	ATP-DEPENDENT CLP PROT	3.75e+01
23	46	62.2	207	1	CLPP_ECOLI	ATP-DEPENDENT CLP PROT	3.75e+01

RESULT	ID	DMPN_PESP	STANDARD	PRT	517 AA
AC	PI9732				
DT	01-FEB-1991 (REL. 17, CREATED)				
DT	01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)				
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)				
DE	PHENOL HYDROXYLASE P3 PROTEIN (EC 1.14.13.7) (PHENOL 2-MONOXYGENASE				
DE	P3 COMPONENT).				
GN	DMPN OR PHENOL.				
OS	PSEUDOMONAS SP. (STRAIN CF600).				
OG	PLASMID PV1150.				
OC	BACTERIA; PROTEOBACTERIA.				
NC	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE: 91072230.				
RA	NORDLUND I., FOWLOWSKI J., SHINGLER V.;				
RT	*Complete nucleotide sequence and polypeptide analysis of				
RT	multicomponent phenol hydroxylase from Pseudomonas sp. strain				
RT	CF600.*				
RL	J. BACTERIOL. 172:6826-6833(1990).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BH;				
RA	TAKAO M., MAEDA Y., OKADA H., MIYAMA K., MORI K., IKE M.;				
RL	FUJITA M.;				
CC	SUBMITTED (MAR-1994) TO EMBL/GENBANK/DDSI DATA BANKS.				
CC	-1- FUNCTION: CATABOLIZES PHENOL, AND SOME OF ITS METHYLATED				
CC	DERIVATIVES. P3 IS REQUIRED FOR GROWTH ON PHENOL, AND FOR				
CC	IN VITRO PHENOL HYDROXYLASE ACTIVITY.				
CC	-1- CATALYTIC ACTIVITY: PHENOL + NADPH + O(2) -> CATECHOL + NADP(+)				
CC	+ H(2)O.				
CC	-1- CORACOR: PAD FLAVOPROTEIN, AND REQUIRES FE(+2) FOR ACTIVITY.				
CC	-1- PATHWAY: FIRST STEP OF PHENOL BIODEGRADATION.				
CC	-1- SUBUNIT: THE MULTICOMPONENT ENZYME PHENOL HYDROXYLASE IS FORMED				
CC	BY P0, P1, P2, P3, P4 AND P5 POLYPEPTIDES.				
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: M60276; G151453; -				
DR	EMBL: D28864; G468469; -				
DR	PIR: D37831; D37831.				

NWSEIPI (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch-pp protein - protein database search, using Smith-Waterman algorithm

on: Thu Sep 2 12:51:38 1999; Maspar time 4.47 Seconds
146,610 Million cell updates/sec

Output not generated.

Title: >US-08-599-226-34
Description: (1-12) from US08599226.pep
Perfect Score: 75
Sequence: 1 ASYLSTASLEY 12

Scoring table: PAM 150
Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: splembl9

1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human
5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.organelle
9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified
13:sp.vertebrate 14:sp.virus

Statistics: Mean 24.013; Variance 31.833; scale 0.754

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	58	77.3	649	3	060167	PROTEIN COMPLEX ASSEMB	7.07e+01
2	57	76.0	503	2	026074	PROTEIN-EXPORT MEMBRAN	1.13e+00
3	53	70.7	897	5	017336	LET 858.	6.94e+00
4	52	69.3	469	2	032204	YVSH PROTEIN.	1.08e+01
5	52	69.3	582	3	074931	ALTERNATIVE NADH-DEHYD	1.08e+01
6	52	69.3	1238	5	061198	FL156.6 PROTEIN.	1.08e+01
7	51	68.0	293	2	051095	CONSERVED HYPOTHETICAL	1.67e+01
8	51	68.0	349	5	017959	MO182.5 PROTEIN.	1.67e+01
9	51	68.0	475	13	093514	AXIAL PROTOCADHERIN (F	1.67e+01
10	51	68.0	932	5	001623	SIMILAR TO LIGAND-GATE	2.56e+01
11	50	66.7	83	9	048385	ORE83.	2.56e+01
12	50	66.7	127	14	011696	NUCLEOPROTEIN (FRAGEN	2.56e+01
13	50	66.7	222	13	P70023	OLFACTOR RECEPTOR (FR	2.56e+01
14	50	66.7	265	5	025144	HROX1.	2.56e+01
15	50	66.7	372	5	P91143	SIMILAR TO ACETYLTRANS	2.56e+01
16	50	66.7	1186	14	055767	PUTATIVE TYROSINE PROT	2.56e+01
17	49	65.3	122	10	P93302	ORF122A.	3.91e+01
18	49	65.3	383	2	Q56675	MANNOSE-SENSITIVE HEMA	3.91e+01
19	49	65.3	910	3	059796	PUTATIVE RECEPTOR ASSO	3.91e+01
20	49	65.3	939	4	015071	KIAA0365 (FRAGMENT).	3.91e+01

21	49	65.3	949	4	060369	KIAA0365 (FRAGMENT).	3.91e+01
22	49	65.3	1406	4	015082	KIAA0377.	3.91e+01
23	48	64.0	290	14	098138	GENE 27 PROTEIN.	5.93e+01
24	48	64.0	311	13	099162	MEANOWA RECEPTOR PROT	5.93e+01
25	48	64.0	343	14	040927	ORE 27.	5.93e+01
26	48	64.0	347	2	P95504	PCB1, PCB2 AND PCB3 GE	5.93e+01
27	48	64.0	350	2	P95503	CHLOROPHYLL A/B ANTENN	5.93e+01
28	48	64.0	357	1	P73958	PRIMASE SMALL SUBUNIT	5.93e+01
29	48	64.0	377	2	084397	HYPOTHEICAL 41.4 KD P	5.93e+01
30	48	64.0	392	2	050873	CONSERVED HYPOTHETICAL	5.93e+01
31	48	64.0	409	2	025853	NADH-BIQUINONE OXIDOR	5.93e+01
32	48	64.0	444	5	P91141	SIMILAR TO ACETYLTRANS	5.93e+01
33	48	64.0	471	10	023552	HYPOTHETICAL 52.8 KD P	5.93e+01
34	48	64.0	511	2	Q43981	PHENOLHYDROXYLASE COMP	5.93e+01
35	48	64.0	626	2	Q46732	COLICIN PROTEIN.	5.93e+01
36	48	64.0	626	2	Q46738	COLICIN PROTEIN.	5.93e+01
37	48	64.0	626	2	Q46740	COLICIN PROTEIN.	5.93e+01
38	48	64.0	626	2	Q46740	COLICIN PROTEIN.	5.93e+01
39	48	64.0	771	3	Q42812	PROLYL DIPEPTIDYL PEPT	5.93e+01
40	48	64.0	989	11	063052	APOLIPOPROTEIN B (FRAG	5.93e+01
41	48	64.0	1026	4	008174	PROTOCADHERIN 42 PRECU	5.93e+01
42	48	64.0	1035	13	057537	NF-PROTOCADHERIN.	5.93e+01
43	48	64.0	1069	11	088185	BH-PROTOCADHERIN-A.	5.93e+01
44	48	64.0	1069	4	060245	PCDH7 (BH-PCDH)A.	5.93e+01
45	48	64.0	1200	4	060247	PCDH7 (BH-PCDH)C.	5.93e+01

ALIGNMENTS

RESULT	ID	PRELIMINARY;	PRT;	649 AA.
AC	060167;			
AD	060167;			
DT	01-AUG-1998 (TREMBL.REL. 07, CREATED)			
DT	01-AUG-1998 (TREMBL.REL. 07, LAST SEQUENCE UPDATE)			
DT	01-AUG-1998 (TREMBL.REL. 07, LAST ANNOTATION UPDATE)			
DE	PROTEIN COMPLEX ASSEMBLY PROTEIN.			
GN	SPBC19F8.03C.			
OS	SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).			
OC	EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;			
OC	SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;			
OC	SCHIZOSACCHAROMYCES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRATN-972H-;			
RA	BECK A., REINHARDT R., WOOD V., RAJANDREAM M.A., BARRELL B.G.;			
RL	SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL; AL023594; E1293401; -			
SO	SEQUENCE 649 AA; 72985 MW; 9C207DB2 CRC32;			

Query Match 77.3%; Score 58; DB 3; Length 649;
Best Local Similarity 72.7%; Pred. No. 7.07e+01;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db	234	NYLSTASLEY 244
Qy	2	SYLSTASLEY 12
RESULT	2	
AC	026074	PRELIMINARY; PRT; 503 AA.
AD	026074;	
DT	01-JAN-1998 (TREMBL.REL. 05, CREATED)	
DT	01-JAN-1998 (TREMBL.REL. 05, LAST SEQUENCE UPDATE)	
DT	01-JAN-1998 (TREMBL.REL. 08, LAST ANNOTATION UPDATE)	
DE	PROTEIN-EXPORT MEMBRANE PROTEIN (SECD).	
GN	HP1550.	
OS	HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).	
OC	BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; HELICOBACTER GROUP;	
OC	HELICOBACTER.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRATN-26695;	
RX	MEDLINE; 97394467.	

RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,
 RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
 RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
 RA LOTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLADEK A.,
 RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
 RA BEIG D.E., GOCAYNE J.D., UTERBACK T.R., PETERSON J.D., KELLEY J.M.,
 RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATTHEY L., WALLIN E.,
 RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
 RA VENTER J.C.;
 RA "The complete genome sequence of the gastric pathogen *Helicobacter*
 RT *pylori* (published erratum appears in Nature 1997 Sep
 RT 25:389(6649):412).";
 RL NATURE 368:539-547(1997).
 DR EMBL; AB000652; G2314730; -;
 DR TIGR; HPI550; -;
 KW HYPOTHETICAL PROTEIN.
 SO SEQUENCE 503 AA; 54247 MW; 9A76592C CRC32;

Query Match 76.0%; Score 57; DB 2; Length 503;
 Best Local Similarity 80.0%; Pred. No. 1.13e+00;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

38 YLSTASLEY 47
 ||| ||| |||
 QY 3 YLSTASLEY 12

RESULT 3 PRELIMINARY; PRT; 897 AA.
 AC 017336;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
 DE LET 858.
 GN LET-858.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
 CC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-N2 (BRISTOL);
 RA KELLY W.G., COLES L.H., FIRE A.Z.;
 RL GENETICS 0:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MATTHEWS L.;
 RL SUBMITTED (JUN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; U19615; G987227; -;
 DR EMBL; Z81525; E1351651; -;
 SO SEQUENCE 897 AA; 104268 MW; E1E3EA36 CRC32;

Query Match 70.7%; Score 53; DB 5; Length 897;
 Best Local Similarity 60.0%; Pred. No. 6.94e+00;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 491 YLTMOSSUDY 500
 ||: ||: ||: ||:
 QY 3 YLSTASLEY 12

RESULT 4 PRELIMINARY; PRT; 469 AA.
 AC 032204;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE YVSH. PROTEIN.
 GN YVSH.
 OS BACILLUS SUBTILIS.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
 CC BACILLUS.
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-168;
 RA MEDLINE; 98044033.
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
 RA AZEVEDO V., BERTERO M.G., BESSIERES P., POLOTIN A., BORCHERT S.,
 RA BORRIS R., BOUSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
 RA CHOI S.K., CODANI J.M., CONNERTON I.F., CUMINGS N.J., DANIEL R.A.,
 RA DENIZOT F., DEVINE K.J., DUSTERHOFT A., EHRICH S.D., EMERSON P.T.,
 RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,
 RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZI A., GALLERON N.,
 RA GHM S.Y., GLASER P., GOREAU A., COLIGHTLY E.J., GRANDI G.,
 RA GUSEPPI G., GUY B.J., HAGA K., HAICH J., HARWOOD C.R., HENAT A.,
 RA HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,
 RA JORIS B., KAKAMATA D., KASAHARA Y., KLAER-BLANCARD M., KLEIN C.,
 RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROH S., KUANO M.,
 RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,
 RA MEDINA N., MELLO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,
 RA MOONE D., O'REILLY M., OGAMA K., OGIMARA A., OUDGA B., PARK S.H.,
 RA PARRO V., POHL T.M., PORTELELLA D., PORODOLIK S., PRESCOTT A.M.,
 RA PRESECAN E., PUTIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,
 RA RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADAIE Y.,
 RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCORFONE F.,
 RA SEKIGUCHI J., SEKONSKA A., SEROR S.J., SEROR P., SHIN B.S., SOLO B.,
 RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,
 RA TAKEUCHI M., TAMAKOSHI A., TANAKA T., TERSTRA P., TOGNONI A.,
 RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTI A.,
 RA VIARI A., WAMBUET R., WEDLER E., WEDLER H., WEITENGER T.,
 RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUNOTO K., YARA K.,
 RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.,
 RT "The complete genome sequence of the gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RL NATURE 390:249-256(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA WIPAT A., BRIGNELL C.S., GUY J.B., ROSE M., EMERSON P.T.,
 RA HARWOOD C.R.;
 RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; Z99121; E1186022; -;
 DR EMBL; A1223978; E1249790; -;
 DR PFAM; PF00324; aa_permeases; 1.
 SO SEQUENCE 469 AA; 50258 MW; 49186162 CRC32;

Query Match 69.3%; Score 52; DB 2; Length 469;
 Best Local Similarity 45.5%; Pred. No. 1.08e+01;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 376 TFLTTATLAY 386
 :||: ||: ||: ||:
 QY 2 SYLSTASLEY 12

RESULT 5 PRELIMINARY; PRT; 582 AA.
 AC 074931;
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE ALTERNATIVE NADH-DEHYDROGENASE PRECURSOR (EC 1.6.5.3)
 DE (NADH DEHYDROGENASE (UBIQUINONE)) (UBIQUINONE REDUCTASE)
 DE (TYPE I DEHYDROGENASE) (COMPLEX I DEHYDROGENASE).
 GN NDH2.
 OS YARROWIA LIPOLYTICA (CANDIDA LIPOLYTICA).
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
 CC DIPODASCACEAE; YARROWIA.
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-E150;
 RA KERSCHER S.J., BRANDT U.,
 RT "identification of the YLNDH2 Gene Encoding the Alternative
 NADH:Ubiquinone Oxidoreductase from Yarrowia lipolytica."
 RL SUBMITTED (JUN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 DE -1- COFACTOR: FAD; IRON-SULFUR.
 DR EMBL: AJ006852; E1330342; -
 KW SIGNAL; OXIDOREDUCTASE.
 FT SIGNAL 1 79
 FT CHAIN 80 582
 SQ SEQUENCE 582 AA; 65814 MW; 0460C796 CRC32;
 POTENTIAL.

Query Match 69.3%; Score 52; DB 3; Length 582;
 Best Local Similarity 54.5%; Pred. No. 1.08e+01;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 81 TYLSA1NSLGY 91
 :|||: |||
 2 SYLSTASLEY 12

RESULT 6 PRELIMINARY; PRT; 1238 AA.
 ID 061198;
 AC 061198;
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE F15E6.6 PROTEIN.
 GN F15E6.6.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIDA; RHABDITIDA;
 OC RHABDITIDA; RHABDITIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE: 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAKINS T., HILLER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIKKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SWALDON N., SMITH A., SONNHAMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
 "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans."
 GN NATURE 368:32-38(1994).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA MILLER N., STELLYES L., BRADSHAW H., KEPLER D.,
 RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA WATERSTON R.,
 RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF038614; G2702437; -
 DR PROSITE: PS00197; 2FE2S_FERREDOXIN; 1.
 KW IRON-SULFUR.
 SQ SEQUENCE 1238 AA; 135726 MW; 9408BB7C CRC32;

Query Match 69.3%; Score 52; DB 5; Length 1238;
 Best Local Similarity 70.0%; Pred. No. 1.08e+01;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

DB 702 YLETOSSLYN 711
 ||| |||
 QY 3 YLSTASLEY 12

RESULT 7 PRELIMINARY; PRT; 293 AA.
 ID 051095;
 AC 051095;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE CONSERVED HYPOTHETICAL PROTEIN.
 GN BB0068.
 OS BORRELLIA BURGDORFERI (LYME DISEASE SPIROCHETE).
 OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELLIA.
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35210 / B31;
 RX MEDLINE: 98065943.
 RA PRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
 RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
 RA VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
 RA UTERBACK T., WATHEY L., McDONALD L., ARTIACH P., BOWMAN C.,
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
 RA SMITH H.O., VENTER J.C.,
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
 burgdorferi."
 RL NATURE 390:580-586(1997).
 DR EMBL: AE001120; G2687956; -
 DR TIGR: BB0068; -
 SQ SEQUENCE 293 AA; 33278 MW; 3FBB9E2 CRC32;

Query Match 68.0%; Score 51; DB 2; Length 293;
 Best Local Similarity 70.0%; Pred. No. 1.67e+01;
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

DB 198 AYLSTPNSLE 207
 :|||: |||
 QY 2 SYLSTASLE 11

RESULT 8 PRELIMINARY; PRT; 349 AA.
 ID 017959;
 AC 017959;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
 DE M01B2.5 PROTEIN.
 GN M01B2.5.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIDA; RHABDITIDA;
 OC RHABDITIDA; RHABDITIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA LLOYD C.,
 RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE: 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAKINS T., HILLER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIKKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SWALDON N., SMITH A., SONNHAMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
 "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans."
 GN NATURE 368:32-38(1994).
 DR EMBL: Z83116; E1348127; -
 RT SEQUENCE 349 AA; 40017 MW; 90870FE2 CRC32;

Query Match 68.0%; Score 51; DB 5; Length 349;
Best Local Similarity 50.0%; Pred. No. 1.67e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 198 YLENSGSLDY 207
Qy 3 YLSTASSLEY 12

RESULT 9 PRELIMINARY; PRT; 475 AA.

ID 093514
AC 093514;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE AXIAL PROTOCADHERIN (FRAGMENT).
GN AXPC.
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOBATRACHIA; PIPOIDEA; PIPIIDAE; XENOPODINAE; XENOPUS.

[1]
SEQUENCE FROM N.A.
YAMAMOTO A., DEROBERTIS E.M.;
"Xenopus axial protocadherin";
RT SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RL EMBL; AF053469; G3598688; -;
FT NON_TER 1
FT SEQUENCE 475 AA; 52268 MW; 2A681544 CRC32;

Query Match 68.0%; Score 51; DB 13; Length 475;
Best Local Similarity 50.0%; Pred. No. 1.67e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 341 FLQTTSLDY 350
Qy 3 YLSTASSLEY 12

RESULT 10 PRELIMINARY; PRT; 932 AA.

ID 001623
AC 001623;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SIMILAR TO LIGAND-GATED IONIC CHANNEL PROTEINS.
GN ZC196.7.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.

[1]
SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERRS M.,
RA BOWFIELD J., BURTON J., CONNELL M., COPEY T., COOPER J., COULSON A.,
RA GRAYTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFFEN L., ROOFA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULLSTON J.,
RA THERY-MEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans";
RT NATURE 368:32-38(1994).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MURRAY J.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U97007; G1938466; -;
DR PFAM; PF00060; 119-chan; 2;
SO SEQUENCE 932 AA; 106836 MW; 469F217A CRC32;

Query Match 68.0%; Score 51; DB 5; Length 932;
Best Local Similarity 54.5%; Pred. No. 1.67e+01;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 753 AYLMSTISLEY 763
Qy 2 YLSTASSLEY 12

RESULT 11 PRELIMINARY; PRT; 83 AA.

ID 048385
AC 048385;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ORF83.
OS STREPTOCOCCUS THERMOPHILUS BACTERIOPHAGE TP-J34.
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; TAILED PHAGES; SIPHOVIRIDAE.
OC [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TP-J34;
RX MEDLINE; 98122991.

RA NEVE H., ZENZ K.I., DESIERE F., KOCH A., HELLER K.J., BRUSSOW H.;
RT "Comparison of the lysogeny modules from the temperate streptococcus
RT thermophilus bacteriophages TP-J34 and Sf121: implications for the
RT modular theory of phage evolution";
RL VIROLOGY 241:61-72(1998).
DR EMBL; AF020798; G2887100; -;
SO SEQUENCE 83 AA; 9876 MW; 250B3F9C CRC32;

Query Match 66.7%; Score 50; DB 9; Length 83;
Best Local Similarity 60.0%; Pred. No. 2.56e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 71 YLETSAFLEY 80
Qy 3 YLSTASSLEY 12

RESULT 12 PRELIMINARY; PRT; 127 AA.

ID 011696
AC 011696;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE NUCLEOPROTEIN (FRAGMENT).
OS MEASLES VIRUS (SUBACUTE SCLEROSE PANENCEPHALITIS VIRUS).
OC VIRUSES; SSRNA NEGATIVE-STRAND VIRUSES; MONONEGAVIRALES;
OC PARAMYXOVIRIDAE; PARAMYXOVIRINAE; MORBILLIVIRUS.

[1]
SEQUENCE FROM N.A.
RC STRAIN-92-E;
RX MEDLINE; 97278133.
RA YAMAGUCHI S.;
RT "Identification of three lineages of wild measles virus by nucleotide
RT sequence analysis of N, P, M, F, and L genes in Japan";
RL J. MED. VIROL. 52:113-120(1997).
DR EMBL; D87487; D1020995; -;
KW NUCLEOPROTEIN.
FT NON_TER 1
FT SEQUENCE 127 AA; 13950 MW; 42D75A2C CRC32;

Query Match 66.7%; Score 50; DB 14; Length 127;
Best Local Similarity 45.5%; Pred. No. 2.56e+01;
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Db 66 AAYLPSTPLD 76
 1 ASYLSTASSLEY 11

RESULT 13
 ID P70023 PRELIMINARY; PRT: 222 AA.
 AC P70023;
 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
 DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE OLFACORY RECEPTOR (FRAGMENT).
 OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
 NC MESOBATRACHIA; PIPOIDEA; PIPOIDAE; XENOPODINAE; XENOPUS.
 RP SEQUENCE FROM N.A.
 RA FREITAG J.;
 [2]
 SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 [2]
 SEQUENCE FROM N.A.
 RX MEDLINE; 96112032.
 RA FREITAG J., KRIGER J., STROTMAN J., BREER H.;
 RT "Two classes of olfactory receptors in *Xenopus laevis*."
 RL NEURON 15:1383-1392(1995).
 DR EMBL: Y08352; E273933;
 DR PFAM: PF00001; 7tm_1, 1.
 FT NON_TER 1 1
 FT NON_TER 222 222
 SQ SEQUENCE 222 AA; 24658 MW; F18F14F8 CRC32;

Query Match 66.7%; Score 50; DB 13; Length 222;
 Best Local Similarity 70.0%; Pred. No. 2.56e+01;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 96 YASTISSLEY 105
 3 YLSTASSLEY 12

RESULT 14
 ID Q25144 PRELIMINARY; PRT: 265 AA.
 AC Q25144;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 RT HROX1.
 NC HALIOTIS RUFESCENS (CALIFORNIA RED ABALONE).
 OC EUKARYOTA; METAZOA; MOLLUSCA; GASTROPODA; PROSOBRANCHIA;
 NC ARCHAGOGASTROPODA; HALIOTIDAE; HALIOTIS.
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 97388321.
 RA DEGNAN B.M., DEGNAN S.M., FENTENANT G., MORSE D.E.;
 RT "A *mx* homeobox gene in the gastropod mollusc *Haliotis rufescens* is
 RT differentially expressed during larval morphogenesis and
 RT metamorphosis."
 RL FEBS LETT. 411:119-122(1997).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL: X75217; GA07415; -.
 DR PROSITE: PS00027; HOMEBOX_1, 1.
 DR PFAM: PF00046; homeobox; 1.
 KW HOMEBOX; DNA-BINDING; NUCLEAR PROTEIN.
 SQ SEQUENCE 265 AA; 29579 MW; B68A753D CRC32;

Query Match 66.7%; Score 50; DB 5; Length 265;
 Best Local Similarity 58.3%; Pred. No. 2.56e+01;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Db 81 GSYLSMSSSKDY 92
 :||||:|:

QY 1 ASYLSTASSLEY 12

RESULT 15
 ID P91143 PRELIMINARY; PRT: 372 AA.
 AC P91143;
 DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
 DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE SIMILAR TO ACETYLTRANSFERASES.
 CN C37H5.2.
 OS CAENORHABDITIS ELEGANS.
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIDA; RHABDITIDAE;
 NC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE; 94150718.
 RA WILSON R., AINSICOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., CORSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU 2, DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HARKINS T., HILLER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THERIERY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*."
 RL NATURE 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA DAVIDSON S., GILLAM B.;
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA WATERSTON R.;
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U88315; G1825777; -.
 DR PFAM: PF00561; abhydrolase; 1.
 KW TRANSFERASE.
 SQ SEQUENCE 372 AA; 42139 MW; 5214F159 CRC32;

Query Match 66.7%; Score 50; DB 5; Length 372;
 Best Local Similarity 50.0%; Pred. No. 2.56e+01;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Search completed: Thu Sep 2 12:52:11 1999
 Job time : 33 secs.

THIS PAGE BLANK (USPTO)

FT TRANSMEM 54 74 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT DOMAIN 75 1038 (POTENTIAL).
 FT DOMAIN 100 175 EXTRACELLULAR (POTENTIAL).
 FT REPEAT 421 522 WFC 1.
 FT REPEAT 592 668 SR1.
 FT REPEAT 677 754 SR2.
 FT REPEAT 742 804 SR3.
 FT DOMAIN 830 899 WFC 2.
 FT DOMAIN 939 1020 WFC 3.
 FT CARBOHYD 179 179 WFC 4.
 FT CARBOHYD 287 287 POTENTIAL.
 FT CARBOHYD 520 520 POTENTIAL.
 FT CARBOHYD 666 666 POTENTIAL.
 FT CARBOHYD 752 752 POTENTIAL.
 FT CARBOHYD 821 821 POTENTIAL.
 SO SEQUENCE 1038 AA; 115514 MW; DC4DAFF5 CRC32;

Query Match 68.0%; Score 51; DB 1; Length 1038;
 Best Local Similarity 60.0%; Pred. No. 4.04e+00;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 497 YLNTDGLAY 506
 11:111111
 QY 3 YLSTASSLEY 12

RESULT 2
 ID DHSD_BOVIN STANDARD; PRT; 158 AA.
 AC 095123;
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
 DE SUCCINATE DEHYDROGENASE (UBIQUINONE) CYTOCHROME B SMALL SUBUNIT
 DE PRECURSOR (CYB) (SUCCINATE-UBIQUINONE REDUCTASE MEMBRANE ANCHOR
 DE SUBUNIT) (Ops3).
 GN SDHD OR SDH4.
 OS BOS TAURUS (BOVINE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HEART;
 RA SHENOY S.K., YU L., YU C.A.;
 RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- SUBUNIT: COMPOSED OF A 27 KD IRON PROTEIN (IP), A 70 KD
 CC FLAVOPROTEIN (FP) AND A CYTOCHROME B COMPOSED OF TWO INTEGRAL
 CC MEMBRANE PROTEINS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U50987; G1575011; -
 KW TRICARBOXYLIC ACID CYCLE; ELECTRON TRANSPORT; HEME; TRANSMEMBRANE;
 KM MITOCHONDRION; TRANSIT PEPTIDE.
 FT TRANSIT 1 55
 FT CHAIN 56 158
 FT TRANSMEM 70 90
 FT TRANSMEM 125 141 CYTOCHROME B SMALL SUBUNIT.
 FT TRANSMEM 158 AA; 17096 MW; 703D5238 CRC32;
 SO SEQUENCE

Query Match 66.7%; Score 50; DB 1; Length 158;
 Best Local Similarity 41.7%; Pred. No. 6.58e+00;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 81 AAYLPCSAMDY 92
 1111111111
 QY 1 ASYLSTASSLEY 12

RESULT 3
 ID DHSD_HUMAN STANDARD; PRT; 159 AA.
 AC 014521;
 DT 15-DEC-1998 (REL. 37, CREATED)
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
 DE SUCCINATE DEHYDROGENASE (UBIQUINONE) CYTOCHROME B SMALL SUBUNIT
 DE PRECURSOR (CYB) (SUCCINATE-UBIQUINONE REDUCTASE MEMBRANE ANCHOR
 DE SUBUNIT).
 GN SDHD OR SDH4.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE: 98194224.
 RA HIRAWAKE H., TANIMAKI M., KIJIMA S., KITA K.;
 RT "Cytochrome b in human complex II (succinate-ubiquinone
 RT oxidoreductase): cDNA cloning of the components in liver mitochondria
 RT and chromosome assignment of the genes for the large (SDHC) and small
 RT (SDHD) subunits to 1q21 and 11q23."
 RL CYTOGENET. CELL GENET. 79:132-138(1997).
 CC -1- SUBUNIT: COMPOSED OF A 27 KD IRON PROTEIN (IP), A 70 KD
 CC FLAVOPROTEIN (FP) AND A CYTOCHROME B COMPOSED OF TWO INTEGRAL
 CC MEMBRANE PROTEINS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB006202; D1022913; -
 DR MIN: 602690; -
 KW TRICARBOXYLIC ACID CYCLE; ELECTRON TRANSPORT; HEME; TRANSMEMBRANE;
 KM MITOCHONDRION; TRANSIT PEPTIDE.
 FT TRANSIT 1 56
 FT CHAIN 57 159
 FT TRANSMEM 71 91
 FT TRANSMEM 126 142 CYTOCHROME B SMALL SUBUNIT.
 FT TRANSMEM 159 AA; 17043 MW; F4221825 CRC32;
 SO SEQUENCE

Query Match 66.7%; Score 50; DB 1; Length 159;
 Best Local Similarity 41.7%; Pred. No. 6.58e+00;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 82 AAYLPCSAMDY 93
 1111111111
 QY 1 ASYLSTASSLEY 12

RESULT 4
 ID B4AR_MELGA STANDARD; PRT; 428 AA.
 AC P43141;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DE BETA-4C ADRENERGIC RECEPTOR.
 GN ADRB4C.
 OS MELEAGRIS GALLOPAGO (COMMON TURKEY).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
 OC NEOGNATHAE; GALLIFORMES; MELEGRIDIDAE; MELEAGRIS.

RN [1]
 RX SEQUENCE FROM N.A.
 RX MEDLINE: 95014249.
 RA CHEN X.-H., HARDEN T.K., NICHOLAS R.A.:
 RT "Molecular cloning and characterization of a novel beta-adrenergic
 RT receptor."
 RL J. BIOL. CHEM. 269:24810-24819(1994).
 CC -1- FUNCTION: BETA-ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
 CC INDUCED ACTIVATION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
 CC PROTEINS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: BROAD TISSUE DISTRIBUTION.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U13977; G556604; -
 DR EMBL: U13978; G555882; -
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 DR HSSP: P07700; 1DEP.
 KM G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KM MULTIGENE FAMILY; PHOSPHORYLATION; LIPOPROTEIN; PALMITATE.
 FT DOMAIN 1 25
 FT TRANSMEM 26 49
 FT DOMAIN 50 58
 FT TRANSMEM 59 77
 FT TRANSMEM 78 97
 FT TRANSMEM 98 119
 FT DOMAIN 120 141
 FT TRANSMEM 142 164
 FT TRANSMEM 165 189
 FT TRANSMEM 190 211
 FT TRANSMEM 212 261
 FT TRANSMEM 262 283
 FT TRANSMEM 284 294
 FT TRANSMEM 295 315
 FT DOMAIN 316 428
 FT CARBOHYD 8
 FT CARBOHYD 13
 FT DISULFID 96
 FT DISULFID 175
 FT LIPID 329
 FT LIPID 329
 FT SEQUENCE 428 AA; 47398 MW; 88794F0C CRC32;
 Query Match 66.7%; Score 50; DB 1; Length 428;
 Best Local Similarity 40.0%; Pred. No. 6.58e+00;
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 Db 122 YLATABLOY 131
 QY 3 YLATABLOY 12
 RESULT 5
 DMPN_PSESP STANDARD; PRT; 517 AA.
 AC P19732;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE PHENOL HYDROXYLASE P3 PROTEIN (EC 1.14.13.7) (PHENOL 2-MONOXYGENASE
 DE P3 COMPONENT).
 DE DMPN OR PHEA4.
 OS PSEUDOMONAS SP. (STRAIN CF600).
 OC PLASMID PVI150.
 OC BACTERIA: PROTEOBACTERIA.
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE: 91072230.
 RA NORDLUND I., POMLOWSKI J., SHINGLER V.:
 RT "Complete nucleotide sequence and polypeptide analysis of
 RT multicomponent phenol hydroxylase from Pseudomonas sp. strain
 RT CF600."
 RL J. BACTERIOL. 172:6826-6833(1990).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BH;
 RA TAKEO M., NAEDA Y., OKADA H., MIYAMA K., MORI K., IKE M.,
 RA FUJITA M.;
 RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: CATABOLIZES PHENOL, AND SOME OF ITS METHYLATED
 CC DERIVATIVES. P3 IS REQUIRED FOR GROWTH ON PHENOL, AND FOR
 CC IN VITRO PHENOL HYDROXYLASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: PHENOL + NADPH + O(2) -> CATECHOL + NADP(+) + H(2)O.
 CC -1- COFACTOR: FAD FLAVOPROTEIN, AND REQUIRES FE(+2) FOR ACTIVITY.
 CC -1- PATHWAY: FIRST STEP OF PHENOL BIODEGRADATION.
 CC -1- SUBUNIT: THE MULTICOMPONENT ENZYME PHENOL HYDROXYLASE IS FORMED
 CC BY P0, P1, P2, P3, P4 AND P5 POLYPEPTIDES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M60276; G151453; -
 DR EMBL: D28864; G468469; -
 DR PIR: D37831; D37831.
 KM AROMATIC HYDROCARBONS CATABOLISM; OXIDOREDUCTASE; MONOOXYGENASE;
 KM FLAVOPROTEIN; FAD; IRON; PLASMID.
 SQ SEQUENCE 517 AA; 60522 MW; FB61602A CRC32;
 Query Match 66.7%; Score 50; DB 1; Length 517;
 Best Local Similarity 50.0%; Pred. No. 6.58e+00;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 Db 101 FLAVASPLEY 110
 QY 3 YLATABLOY 12
 RESULT 6
 ID 5HT1_DROME STANDARD; PRT; 564 AA.
 AC P20905;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE 5-HYDROXYTRYPTAMINE RECEPTOR 1 (5-HT RECEPTOR) (SEROTONIN RECEPTOR).
 GN 5HT-R1 OR 5-HT7.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
 OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
 OC DROSOPHILIDAE; DROSOPHILA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OREGON-R; TISSUE-HEAD;
 RX MEDLINE: 91062395.
 RA WITZ P., AMLAIRY N., PLASSAT J.-L., MAROTEAUX L., BORRELLI E., HEN R.;
 RT "Cloning and characterization of a Drosophila serotonin receptor that
 RT activates adenylylate cyclase."
 RT PROC. NATL. ACAD. SCI. U.S.A. 87:8940-8944(1990).
 CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR
 CC 5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTION
 CC AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE
 CC CYCLASE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: HEAD.

```
CC -! SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC BUT WITH ONE EXTRA POTENTIAL TRANSMEMBRANE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M55533; GI56725; -.
DR PIR, A38271; A38271.
DR GCRDB, GCR_0023; -.
DR FLYBASE, FBgn0004573; 5-HT7.
DR PROSITE, PS00237; G_PROTEIN_RECEPTOR, 1.
DR PFAM, PF00001; 7tm_t, 1.
KM G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT.
FT TRANSMEM 29 51
FT TRANSMEM 165 188
FT TRANSMEM 189 198
FT TRANSMEM 199 222
FT TRANSMEM 223 236
FT TRANSMEM 227 258
FT TRANSMEM 259 278
FT TRANSMEM 279 302
FT TRANSMEM 303 330
FT TRANSMEM 331 353
FT TRANSMEM 354 454
FT TRANSMEM 455 476
FT TRANSMEM 477 487
FT TRANSMEM 488 510
FT TRANSMEM 511 564
FT TRANSMEM 564 595
FT TRANSMEM 595 608
FT TRANSMEM 608 614
FT TRANSMEM 614 614
SQ SEQUENCE 564 AA; 60861 MW; 312369E8 CRC32;
Query Match 66.7%; Score 50; DB 1; Length 564;
Best Local Similarity 50.0%; Pred. No. 6,58e+00;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Db 261 YLAITKPLEY 270
Qy 3 YLTASSLEY 12
|:::||||
-----
RESULT 7
ID YDH3_SCHPO STANDARD: PRT: 352 AA.
OC 092348;
OC 01-NOV-1997 (REL. 35, CREATED)
OC 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
OC 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 39.7 KD PROTEIN C6G9.03C IN CHROMOSOME I.
GN SPAC669.03C.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., CONNOR R.E.;
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DDJJ DATA BANKS.
CC -! SIMILARITY: SOME. TO YEAST YNL206C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

DB	QY	303 SYLNASSLE 312	65.3%	Score 49:	DB 1:	Length 352:
Query Match	Best Local Similarity	80.0%	Pred. No. 1.06e+01:	Mismatches	1:	Indels 0:
Matches	8:	Conservative	1:	Mismatches	1:	Indels 0:
DB	303 SYLNASSLE 312	65.3% <td>Score 49:<td>DB 1:<td>Length 352:</td><td></td></td></td>	Score 49: <td>DB 1:<td>Length 352:</td><td></td></td>	DB 1: <td>Length 352:</td> <td></td>	Length 352:	
QY	2 SYLNASSLE 11	80.0% <td>Pred. No. 1.06e+01:</td> <td>Mismatches</td> <td>1:</td> <td>Indels 0:</td>	Pred. No. 1.06e+01:	Mismatches	1:	Indels 0:
RESULT	8	STANDARD:	PRT:	441 AA.		
ID	SECY_MYCTU	STANDARD:	PRT:	441 AA.		
AC	P94926;					
DT	15-JUL-1998 (REL. 36, CREATED)					
DT	15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)					
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)					
DE	PREPROTEIN TRANSLOCASE SECY SUBUNIT.					
GN	SECY OR RV0732 OR MTV041.06.					
OS	MYCOBACTERIUM TUBERCULOSIS, AND MYCOBACTERIUM BOVIS.					
OC	BACTERIA: FIRMICUTES: ACTINOBACTERIA: ACTINOBACTERIACEAE:					
CC	ACTINOMYCETALES: CORYNEBACTERIALE: MYCOBACTERIACEAE: MYCOBACTERIUM.					
FN	[1]					
RP	SEQUENCE FROM N.A.					
RC	SPECIES=M.TUBERCULOSIS; STRAIN=H37RV;					
RA	MEDLINE: 98295987.					
RA	COLE S.T., BOSCH R., PARKHILL J., GARNIER T., CHURCHER C., HARRIS D.,					
RA	GORDON S.V., EIGMEYER K., GAS S., BARRY C.E., ILLI, TEKATA F.,					
RA	BADCOCK K., BASHAM D., BROWN D., CHILLINGWORTH T., CONNOR R.,					
RA	DAVIES R., DEVILIN K., FELTWELL T., GENTLES S., HAWLIN N., HOLROYD S.,					
RA	HONNBER T., JAGELS K., KROGH A., MCLEAN J., MOULE S., MURPHY T.,					
RA	OLIVER S., OSBORNE J., QUAIL M.A., RAJANDREAN M.A., ROGERS J.,					
RA	RUTTER S., SEEGER K., SKELTON S., SQUARES S., SQUARES R., SULSTON J.E.,					
RA	TAYLOR K., WHITEHEAD S., BARRELL B.G.;					
RT	"Deciphering the biology of Mycobacterium tuberculosis from the					
RT	complete genome sequence."					
RL	NATURE 393:537-544(1998).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	SPECIES=M.BOVIS; STRAIN=BCG;					
RA	KIM J.K., CHOE Y.K.;					
RL	SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.					
CC	-1- FUNCTION: INVOLVED IN PROTEIN EXPORT. INTERACTS WITH SECA AND SECE					
CC	TO ALLOW THE TRANSLOCATION OF PROTEINS ACROSS THE PLASMA MEMBRANE,					
CC	BY FORMING PART OF A CHANNEL. (BY SIMILARITY).					
CC	-1- SUBUNIT: ONE OF SEVEN SECRETORY PROTEINS (SECA-F & SECY) THAT					
CC	COMPRISE THE PROKARYOTIC PROTEIN TRANSLOCATION APPARATUS					
CC	(BY SIMILARITY).					
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.					
CC	-1- SIMILARITY: BELONGS TO THE SECY/SEC61-ALPHA FAMILY.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.1sb-sib.ch/announce/					
CC	or send an email to license@1sb-sib.ch).					
CC	-----					
CC	EMBL; AL021958; E1253270; .					
DR	EMBL; U77912; G1813627; .					
DR	PROSITE; PS00755; SECY_1; 1.					
DR	PROSITE; PS00756; SECY_2; 1.					
DR	PFAM; PF00344; secy; 1					
FT	PROTEIN TRANSPORT; TRANSMEMBRANE; TRANSLOCATION.					
FT	TRANSMEM 18					
FT	TRANSMEM 38					
FT	TRANSMEM 57					
FT	TRANSMEM 77					
FT	TRANSMEM 78					
FT	TRANSMEM 98					
FT	TRANSMEM 124					
FT	TRANSMEM 144					
FT	TRANSMEM 157					
FT	TRANSMEM 177					
FT	TRANSMEM 180					
FT	TRANSMEM 200					
FT	POTENTIAL.					

FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 272 292 POTENTIAL.
 FT TRANSMEM 318 338 POTENTIAL.
 FT TRANSMEM 382 402 POTENTIAL.
 FT TRANSMEM 404 424 POTENTIAL.
 SQ SEQUENCE 441 AA; 47611 MW; 261FAC6 CRC32;

Query Match 65.3%; Score 49; DB 1; Length 441;
 Best Local Similarity 50.0%; Pred. No. 1.06e+01;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 309 GTYSDPSNLVY 320
 :|||:|:|
 QY 1 ASYLSTASSLEY 12

RESULT 9
 ID YMG1_YEAST STANDARD; PRT; 475 AA.
 AC 003652;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DE HYPOTHEICAL 55.3 KD PROTEIN IN RARI-SCU1 INTERGENIC REGION.
 GN YMR211W OR YMR261.05.
 OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 CC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCETES; SACCCHAROMYCETALES;
 CC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA DEDMAN K., BROWN D., BOWMAN S., BARRELL B.G., RAJANDREAN M.A.,
 RA MALSH S.V.;
 RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial/
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC
 DR EMBL; 249809; G854463;
 DR PFM; PF00091; tubulin; 1.
 KW HYPOTHEICAL PROTEIN.
 SQ SEQUENCE 475 AA; 55312 MW; 18FA4F03 CRC32;

Query Match 65.3%; Score 49; DB 1; Length 475;
 Best Local Similarity 60.0%; Pred. No. 1.06e+01;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 321 YLTATITLGY 330
 :|||:|:|
 QY 3 YLSTASSLEY 12

RESULT 10
 ID GLRK_LYMT STANDARD; PRT; 917 AA.
 AC P26591;
 DT 01-AUG-1992 (REL. 23, CREATED)
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
 DE GLUTAMATE RECEPTOR PRECURSOR.
 OS LYMAEA STAGNALIS (GREAT POND SNAIL).
 CC EUKARYOTA; METAZOA; MOLUSCA; GASTROPODA; PULMONATA; BASOMATOPHORA;
 CC LYMAEIDAE; LYMAEA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE; 92070466.
 RA HUTTON M.L., HAREY R.J., BARNARD E.A., DARLISON M.G.;
 RT "Cloning of a cDNA that encodes an invertebrate glutamate receptor
 subunit.";
 RL FEBS LETT. 292:111-114(1991).

CC -1- FUNCTION: L-GLUTAMATE ACTS AS AN EXCITATORY NEUROTRANSMITTER AT
 MANY SYNAPSES IN THE CENTRAL NERVOUS SYSTEM. THE POSTSYNAPTIC
 ACTIONS OF GLU ARE MEDIATED BY A VARIETY OF RECEPTORS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial/
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC
 DR EMBL; X60086; G9629;
 DR PIR; S15681; ACBAE.
 DR PIR; S18443; S18443.
 DR PFM; PF00060; 1lg_chan; 1.
 KW RECEPTOR; POSTSYNAPTIC MEMBRANE; IONIC CHANNEL; GLYCOPROTEIN; SIGNAL;
 RN TRANSMEMBRANE.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 917 GLUTAMATE RECEPTOR.
 FT DOMAIN 21 558 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 559 578 POTENTIAL.
 FT TRANSMEM 599 617 POTENTIAL.
 FT TRANSMEM 628 646 POTENTIAL.
 FT TRANSMEM 819 839 POTENTIAL.
 FT CARBOHYD 62 62 POTENTIAL.
 FT CARBOHYD 95 95 POTENTIAL.
 FT CARBOHYD 121 121 POTENTIAL.
 FT CARBOHYD 125 125 POTENTIAL.
 FT CARBOHYD 229 229 POTENTIAL.
 FT CARBOHYD 251 251 POTENTIAL.
 FT CARBOHYD 261 261 POTENTIAL.
 FT CARBOHYD 272 272 POTENTIAL.
 FT CARBOHYD 418 418 POTENTIAL.
 FT CARBOHYD 419 419 POTENTIAL.
 FT CARBOHYD 424 424 POTENTIAL.
 FT CARBOHYD 491 491 POTENTIAL.
 FT CARBOHYD 775 775 POTENTIAL.
 FT CARBOHYD 881 881 POTENTIAL.
 SQ SEQUENCE 917 AA; 103139 MW; 879CEBDC CRC32;

Query Match 65.3%; Score 49; DB 1; Length 917;
 Best Local Similarity 36.4%; Pred. No. 1.06e+01;
 Matches 4; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 729 AYLFESTIDY 739
 :||:|:|:|
 QY 2 YLSTASSLEY 12

RESULT 11
 ID YIGL_HAEIN STANDARD; PRT; 272 AA.
 AC P44771;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DE HYPOTHEICAL PROTEIN HI0597.
 GN HI0597.
 OS HAEMOPHILUS INFLUENZAE.
 CC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURILLACEAE;
 CC HAEMOPHILUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RD / KW20;
 RX MEDLINE; 95350630.
 RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
 RA KERLAVAGE A.R., BULL C.J., TOMB J.F., DOUGHERTY B.A., MERRICK J.M.,
 RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
 RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODER A., KELLEY J.M.,
 RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLUM E., COTTON M.D.,
 RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,

CC RA FINE L.D., FITCHMAN J.L., FUHRMANN J.L., GEORGACAN N.S.M.,
CC RA GNHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
CC RA VENTER J.C.,
CC RT "Whole-genome random sequencing and assembly of Haemophilus
CC RT influenzae Rd".
CC RL SCIENCE 269:496-512(1995).
CC CC -1- SIMILARITY: BELONGS TO THE COF/YBHA/YIDA/YIGL (E.COLI) / YOSE/YXEH
CC CC (B.SUBTILIS) FAMILY. STRONG, TO E.COLI YIGL.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL: U32741; G1573586; -.
CC DR TIGR: H10597; -.
CC RR PROSITE; PS01228; COF_1; 1.
CC RR PROSITE; PS01229; COF_2; 1.
CC RR PFWAM; PFW0592; DUF3; 1.
CC RR HYPOTHETICAL PROTEIN.
CC RM SEQUENCE 272 AA: 30523 MW: 753365C CRC32:
CC SQ
CC
CC Query Match 64.0%; Score 48; DB 1; Length 272;
CC Best Local Similarity 50.0%; Pred. No. 1,70e+01;
CC Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
CC
CC Db 261 ARYLTOFGDLY 272
CC | | | | : | : |
CC Qy 1 ASYLSTASSLEY 12
CC
CC
CC RESULT 12
CC ID UPO1.MAIZE STANDARD: PRT: 471 AA.
CC AC P16166;
CC DT 01-APR-1990 (REL. 14, CREATED)
CC DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
CC DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
CC DE FLAVONOL 3-O-GLUCOSYLTRANSFERASE (EC 2.4.1.91) (UDP-GLUCOSE FLAVONOID
CC DE 3-O-GLUCOSYLTRANSFERASE) (BRONZE-1) (BZ-MCC ALLELE).
CC GN BZ1 OR UGT71A1.
CC OS ZEA MAYS (MAIZE).
CC OC EUARIOTIA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
CC OC EUDYLIPTOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES;
CC OC POACEAE; ZEA.
CC [1]
CC [1] SEQUENCE FROM N.A.
CC RA FURTER D., SCHIEFELBEIN J.W., JOHNSTON F., NELSON O.E. JR.;
CC RA "Sequence comparisons of 3 wild-type bronze-1 alleles from zea mays";
CC RA PLANT MOL. BIOL. 11:473-481(1988).
CC [12]
CC RA RAUSTON E.J., ENGLISH J.J., DOONER H.K.;
CC RA "Sequence of three bronze alleles of maize and correlation with the
CC RA genetic fine structure."
CC RT GENETICS 119:185-197(1988).
CC RL -1- FUNCTION: IN THE PRESENCE OF OTHER NECESSARY COLOR FACTORS, THIS
CC CC GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
CC CC PIGMENTS.
CC CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + A FLAVONOL = UDP + FLAVONOL
CC CC 3-O-D-GLUCOSIDE.
CC CC -1- PATHWAY: ONE OF THE ESSENTIAL AND TERMINAL STEPS IN THE
CC CC ANTHOCYANIN BIOSYNTHETIC PATHWAY
CC CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial

```

CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  DR EMBL; X13500; G1030071; -.
CC  DR EMBL; X07940; G22205; -.
CC  DR PIR; S01052; S01052.
CC  DR PIR; S08324; S08324.
CC  DR MAZEDB; 13885; -.
CC  DR PROSITE; PS00375; UPPGT; 1.
CC  DR PFAM; PF00201; UPPGT; 2.
CC  KW TRANSFERASE; GLYCOSYLTRANSFERASE.
CC  SQ SEQUENCE 471 AA; 48769 MM; 8AE03FD2 CRC32;
CC
CC  Query Match 64.0%; Score 48; DB 1; Length 471;
CC  Best Local Similarity 88.9%; Pred. No. 1.70e+01;
CC  Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
DB  47 SFLSTASSL 55
QY  2 SYLSTASSL 10

```

```

RESULT 14
ID UFG3_MAIZE STANDARD; PRT; 471 AA.
AC P16157;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE FLAVONOL 3-O-GLUCOSYLTRANSFERASE (EC 2.4.1.91) (UDP-GLUCOSE FLAVONOID
DE 3-O-GLUCOSYLTRANSFERASE) (BRONZE-1) (Bz-W22 ALLELE).
GN BZ1 OR UGT71A1
OS ZEA MAYS (MAIZE).
OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLOPHYTES; SPERMATOPHYTES; MAGNOLIOPHYTES; LILIOPSIDA; POALES;
OC FOAGACEAE; ZEA.
[1]
RN SEQUENCE FROM N.A.
RP FURTER D., SCHIEFELBEIN J.W., JOHNSTON F., NELSON O.E. JR.;
RT "Sequence comparisons of 3 wild-type bronze-1 alleles from Zea mays.";
RL PLANT MOL. BIOL. 11:473-481(1988).
[2]
RX MEDLINE: 88284304.
RA RALSTON E.J., ENGLISH J.J., DOONER H.K.;
RT "Sequence of three bronze alleles of maize and correlation with the
RT genetic fine structure.";
RL GENETICS 119:185-197(1988).
CC -1- FUNCTION: IN THE PRESENCE OF OTHER NECESSARY COLOR FACTORS, THIS
CC GLYCOSYLATION REACTION ALLOWS THE ACCUMULATION OF ANTHOCYANIN
CC PIGMENTS.
CC -1- CATALYTIC ACTIVITY: UDP-GLUCOSE + A FLAVONOL = UDP + FLAVONOL
CC 3-O-D-GLUCOSIDE.
CC -1- PATHWAY: ONE OF THE ESSENTIAL AND TERMINAL STEPS IN THE
CC ANTHOCYANIN BIOSYNTHETIC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE UDP-GLYCOSYLTRANSFERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X13502; G22506; -
DR EMBL: X07937; G22210; -
DR PIR: S01037; S01037.
DR PIR: S08326; S08326.
MAIZEIDB: 13885; -
PROSITE: PS00375; UDPGT: 1.
PRAM: PF00201; UDPGT: 2.
KM TRANSFERASE: GLYCOSYLTRANSFERASE.
SQ SEQUENCE 471 AA; 48673 MW; 4A3C6193 CRC32;

Query Match 64.0%; Score 48; DB 1; Length 471;
Best Local Similarity 88.9%; Pred. NO. 1.70e+01;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 47 SYLSTASSL 55
QY 2 SYLSTASSL 10

```

```

OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87008385.
RA MANKOVICH J.A., HSU C.-H., KONISKY J.;
RT "DNA and amino acid sequence analysis of structural and immunity
RT genes of colicins Ia and Ib.";
RL J. BACTERIOL. 168:228-236(1986).
[2]
RN SEQUENCE OF 1-40 FROM N.A.
RP MEDLINE: 84264487.
RA MANKOVICH J.A., LAI P.-H., GOKUL N., KONISKY J.;
RT "Organization of the colicin Ib gene. Promoter structure and immunity
RT domain.";
RL J. BIOL. CHEM. 259:8764-8768(1984).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE: 85014128.
RA VARLEY J.M., BOULINOIS G.J.;
RT "Analysis of a cloned colicin Ib gene: complete nucleotide sequence
RT and implications for regulation of expression.";
RL NUCLEIC ACIDS RES. 12:6727-6739(1984).
[4]
RN ERRATUM.
RA VARLEY J.M., BOULINOIS G.J.;
RL NUCLEIC ACIDS RES. 12:8748-8748(1984).
CC -1- FUNCTION: THIS COLICIN IS A CHANNEL-FORMING COLICIN. THIS CLASS OF
CC TRANSMEMBRANE TOXINS DEPOLARIZE THE CYTOPLASMIC MEMBRANE, LEADING
CC TO DISSIPATION OF CELLULAR ENERGY.
CC -1- FUNCTION: COLICINS ARE POLYPEPTIDE TOXINS PRODUCED BY AND ACTIVE
CC AGAINST, ESCHERICHIA COLI AND CLOSELY RELATED BACTERIA.
CC -1- SIMILARITY: BELONGS TO THE CHANNEL FORMING COLICIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: K02071; G144640; -
DR EMBL: X01009; G41142; -
DR EMBL: M13820; G144648; -
DR PIR: A03503; IRECB.
DR PIR: A22503; A22503.
DR PIR: D25035; D25035.
DR PROSITE: PS00276; CHANNEL_COLICIN; 1.
DR PRAM: PF01024; COLICIN; 1.
DR HSSP: P06716; ICIT
KM PLASMID: BACTERIOCIN; COLICIN; TOXIN; TRANSMEMBRANE.
FT TRANSMEM 588 612 POTENTIAL.
SQ SEQUENCE 626 AA; 69923 MW; 983D11B0 CRC32;

Query Match 64.0%; Score 48; DB 1; Length 626;
Best Local Similarity 54.5%; Pred. NO. 1.70e+01;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 333 NYLTHSGLDY 343
QY 2 SYLSTASSLEY 12

```

Search completed: Thu Sep 2 12:51:21 1999
Job time : 9 secs.

THIS PAGE BLANK (USPTO)

TITLE protein-export membrane protein - Helicobacter pylori (strain J99)
ORGANISM #formal_name Helicobacter pylori
#strain J99
DATE 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 05-Mar-1999
ACCESSIONS D71805
REFERENCE A71800
#authors Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Noonan, B.; Guild, B.C.; deJonghe, B.L.; Carmel, G.; Tummalo, P.J.; Caruso, A.; Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Trust, T.J.
#journal Nature (1999) 397:176-180
#title Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.
#cross-references MUID:99120557
#accession D71805
#status Preliminary
#molecule_type DNA
#residues 1-526 #label ARN
#cross-references GB:AE001567; GB:AE001439; NID:g4156065; PID:g4156069
#experimental_source strain J99
GENETICS
#gene secD
CLASSIFICATION #superfamily protein export membrane protein secD
SUMMARY #length 526 #molecular-weight 56796 #checksum 5813
Query Match 76.0%; Score 57; DB 2; Length 526;
Best Local Similarity 80.0%; Pred. No. 6.18e+01;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 60 YLSTASLEY 69
||| ||| |||
OY 3 YLSTASLEY 12
RESULT 3
ENTRY D70048
TITLE ABC transporter (amino acid permease) homolog yvsh - Bacillus subtilis
ORGANISM #type complete
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1998
ACCESSIONS D70048
REFERENCE A69580
#authors Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Allioni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans, A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devine, K.M.; Diesterhoef, A.; Enlrich, S.D.; Emerson, P.T.; Ertlan, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Galleron, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.; Giuseppe, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.; Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi, Y.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Porwolik, S.; Prescott, A.M.; Prescan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.; Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, E.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;

Sekowska, A.; Seror, S.J.; Serior, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Tacconi, E.; Takegi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.; Wandut, R.; Wedler, E.; Wedler, H.; Weitzengrter, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
#journal Nature (1997) 390:249-256
#title The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
#cross-references MUID:98044033
#accession D70048
#status Preliminary; nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-469 #label KUN
#cross-references GB:Z99121; GB:AL009126; NID:g2635827; PID:el186022; PID:g2635847
#experimental_source strain 168
GENETICS
#gene yvsh
CLASSIFICATION #superfamily L-lysine transport protein
SUMMARY #length 469 #molecular-weight 50258 #checksum 4200
Query Match 69.3%; Score 52; DB 2; Length 469;
Best Local Similarity 45.5%; Pred. No. 6.57e+00;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Db 376 TELLTAATLAY 386
||| ||| |||
OY 2 YLSTASLEY 12
RESULT 4
ENTRY D70108
TITLE conserved hypothetical protein BB0068 - Lyme disease spirochaete
ORGANISM #type complete
DATE 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 05-Jun-1998
ACCESSIONS D70108
REFERENCE A70100
#authors Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.; Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb, J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.; Kierlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vglt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman, J.; Utterback, T.; Matthey, L.; McDonald, L.; Attiach, P.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.
#journal Nature (1997) 390:580-586
#title Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
#cross-references MUID:98065943
#accession D70108
#status Preliminary; nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-293 #label KLE
#cross-references GB:AE001120; GB:AE000783; NID:g2687951; PID:g2687956; TIGR:BB0068
#experimental_source strain B31
SUMMARY #length 293 #molecular-weight 33278 #checksum 5223
Query Match 68.0%; Score 51; DB 2; Length 293;
Best Local Similarity 70.0%; Pred. No. 1.03e+01;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 198 AYLSTPNSLE 207

QY 2 SYLSTASSLEY 11 :||||:|

RESULT 5 F71918 #type complete
ENTRY hypothetical protein jhp0552 - Helicobacter pylori (strain J99)
TITLE #formal_name Helicobacter pylori
ORGANISM #variety
DATE 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 13-Feb-1999

ACCESSIONS F71918
REFERENCE A71800
#authors Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Noonan, B.; Guild, B.C.; deGongre, B.L.; Carmel, G.; Tummino, P.J.; Caruso, A.; Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, O.; Taylor, D.E.; Vovis, G.F.; Trust, T.J.
#journal Nature (1999) 397:176-180
#title Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.

#cross-references M01D:99120557
#accession F71918
#status Preliminary
#molecule_type DNA
#residues 1-477 #label ARN
#cross-references GB:AE001487; GB:AE001439; NID:g4155086; PID:g4155090
#experimental_source strain J99

GENETICS jhp0552
SUMMARY #length 477 #molecular-weight 54625 #checksum 3264

Query Match 68.0%; Score 51; DB 2; Length 477;
Best Local Similarity 58.3%; Pred. No. 1.03e+01;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 172 AYRSTANLEY 183 |:||:|

QY 1 ASYLSTASSLEY 12

RESULT 6 S69755 #type complete
ENTRY hypothetical protein YDR203w - yeast (Saccharomyces cerevisiae)
TITLE #formal_name Saccharomyces cerevisiae
ORGANISM #variety
DATE 23-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 12-Dec-1997

ACCESSIONS S69755
REFERENCE S52697
#authors Oliver, K.; Harris, D.
#submission submitted to the EMBL Data Library, March 1995
#accession S69755
#molecule_type DNA
#residues 1-105 #label OLI
#cross-references EMBL:Z48784; MIPS:YDR203w

GENETICS
SUMMARY #map_position 4R #length 105 #molecular-weight 11606 #checksum 5265

Query Match 66.7%; Score 50; DB 2; Length 105;
Best Local Similarity 54.5%; Pred. No. 1.62e+01;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 74 SYLSTASSLEY 84 |||||:|

QY 2 SYLSTASSLEY 12

RESULT 7 S38380 #type complete
ENTRY

TITLE Hroxl protein - California red abalone
ORGANISM #formal_name Haliotis rufescens #common_name California red abalone
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Oct-1997

ACCESSIONS S38380
REFERENCE S38380
#authors Degnan, B.M.; Degnan, S.M.; Morse, D.E.
#submission submitted to the EMBL Data Library, September 1993
#description Expression of Hroxl, a gastropod mollusc Hox homeobox gene, is progressively restricted during morphogenesis from trochophore to veliger larval forms.

#accession S38380
#status Preliminary
#molecule_type mRNA
#residues 1-265 #label DEG
#cross-references EMBL:X75217; NID:g407414; PID:g407415
CLASSIFICATION #superfamily unassigned homeobox proteins; homeobox homology DNA binding; homeobox; nucleus; transcription regulation
KEYWORDS
FEATURE 164-220
SUMMARY #domain homeobox homology #label HOX
#length 265 #molecular-weight 28579 #checksum 8996

Query Match 66.7%; Score 50; DB 2; Length 265;
Best Local Similarity 58.3%; Pred. No. 1.62e+01;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 81 GSYLSMSSKDY 92 :||||:|

QY 1 ASYLSTASSLEY 12

RESULT 8 A55044 #type complete
ENTRY beta-4C-adrenergic receptor - turkey
TITLE #formal_name Melagris gallopavo #common_name common turkey
ORGANISM #variety
DATE 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 08-Sep-1997

ACCESSIONS A55044
REFERENCE A55044
#authors Chen, X.; Harden, T.K.; Nicholas, R.A.
#journal J. Biol. Chem. (1994) 269:24810-24819
#title Molecular cloning and characterization of a novel beta-adrenergic receptor.

#accession A55044
#status Preliminary
#molecule_type DNA
#residues 1-428 #label CHE
#cross-references GB:U13978; NID:g555881; PID:g555882

GENETICS
SUMMARY #introns 416/2
CLASSIFICATION #superfamily vertebrate rhodopsin neurotransmitter receptor; transmembrane protein
KEYWORDS
SUMMARY #length 428 #molecular-weight 47398 #checksum 8085

Query Match 66.7%; Score 50; DB 2; Length 428;
Best Local Similarity 40.0%; Pred. No. 1.62e+01;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 122 YLATRAPLOY 131 |:||:|

QY 3 YLSTASSLEY 12

RESULT 9 D37831 #type complete
ENTRY phenol 2-monooxygenase (EC 1.14.13.7) chain P3 - Pseudomonas
TITLE #formal_name Pseudomonas sp.
ORGANISM #formal_name Pseudomonas sp.
DATE 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 31-Oct-1997

ACCESSIONS D37831
REFERENCE A37831

#authors Nordlund, I.; Powlowski, J.; Shingler, V.
#journal J. Bacteriol. (1990) 172:6826-6833
#title Complete nucleotide sequence and polypeptide analysis of multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CF600.
#cross-references MUID:91072230
#accession D37831
#status preliminary
#molecule-type DNA
#residues 1-517 #label NOR
#cross-references GB:M0276; GB:M37764; NID:g151449; PID:g151453
KEYWORDS oxidoreductase
SUMMARY #length 517 #molecular-weight 60522 #checksum 6342

Query Match 66.7%; Score 50; DB 2; Length 517;
Best Local Similarity 50.0%; Pred. No. 1.62e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 101 FITAVSLEY 110
: : : : :
3 YLSTASSLEY 12

RESULT 10
ENTRY A38271 #type complete
TITLE serotonin receptor 7 - fruit fly (*Drosophila melanogaster*)
ALTERNATE_NAMES 5-hydroxytryptamine receptor 7 (5-HTR7)
ORGANISM #formal_name *Drosophila melanogaster*
DATE 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 04-Sep-1998

ACCESSIONS A38271
REFERENCE A38271
#authors Witz, P.; Amalaky, N.; Plassat, J.L.; Maroteaux, L.; Borrelli, E.; Hen, R.
#journal Proc. Natl. Acad. Sci. U.S.A. (1990) 87:8940-8944
#title Cloning and characterization of a *Drosophila* serotonin receptor that activates adenylylate cyclase.
#cross-references MUID:91062395
#accession A38271
#molecule-type mRNA
#residues 1-564 #label WIT
#cross-references GB:M5533
#note The authors translated the codon CTT for residue 213 as Ala, GTG for residue 215 as Leu, ATG for residue 280 as Leu and ATG for residue 551 as Asn

GENETICS
#gene FLYBase:5-HT7
#cross-references FLYBase:FBgn0004573
CLASSIFICATION #superfamily octopamine receptor type I
KEYWORDS G protein-coupled receptor; glycoprotein; membrane protein; phosphoprotein
SUMMARY #length 564 #molecular-weight 60934 #checksum 3324

Query Match 66.7%; Score 50; DB 2; Length 564;
Best Local Similarity 50.0%; Pred. No. 1.62e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 261 YLSTASSLEY 270
: : : : :
3 YLSTASSLEY 12

RESULT 11
ENTRY T03180 #type complete
TITLE tyrosine protein kinase homolog - Chilo iridescent virus
ORGANISM #formal_name Chilo iridescent virus
DATE 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999

ACCESSIONS T03180
REFERENCE 214834
#authors Bahr, U.; Tidona, C.A.; Darai, G.
#journal Virus Genes (1997) 15:235-245
#title The DNA sequence of Chilo iridescent virus between the genome

#accession coordinates 0.101 and 0.391; similarities in coding strategy between insect and vertebrate iridoviruses.
#status T03180
#molecule-type DNA
#residues 1-1186 #label BAH
#cross-references EMBL:AF003534; NID:g2738385; PID:g2738451
SUMMARY #length 1186 #molecular-weight 138020 #checksum 6252

Query Match 66.7%; Score 50; DB 2; Length 1186;
Best Local Similarity 54.5%; Pred. No. 1.62e+01;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 503 TYLSDPSNLY 513
: : : : :
2 YLSTASSLEY 12

RESULT 12
ENTRY G70822 #type complete
TITLE probable secy protein - *Mycobacterium tuberculosis* (strain H37Rv)
ORGANISM #formal_name *Mycobacterium tuberculosis*
DATE 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 21-Nov-1998

ACCESSIONS G70822
REFERENCE A70500
#authors Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eigemeier, K.; Gas, S.; Barry III, C.E.; Tekala, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rayment, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
#journal Nature (1998) 393:537-544
#title Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence.
#cross-references MUID:98295987
#accession G70822
#status preliminary; nucleic acid sequence not shown; translation not shown

GENETICS
#molecule-type DNA
#residues 1-441 #label COL
#cross-references GB:AL021958; GB:AL123456; NID:g3261536; PID:e1253270;
#experimental_source strain H37Rv
CLASSIFICATION #superfamily preprotein translocase secy
SUMMARY #length 441 #molecular-weight 47611 #checksum 3743

Query Match 65.3%; Score 49; DB 2; Length 441;
Best Local Similarity 50.0%; Pred. No. 2.52e+01;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 309 GYLDPSNLY 320
: : : : :
1 ASYLTASSLEY 12

RESULT 13
ENTRY S55093 #type complete
TITLE hypothetical protein YMR211w - yeast (*Saccharomyces cerevisiae*)
ALTERNATE_NAMES hypothetical protein YMR261.05
ORGANISM #formal_name *Saccharomyces cerevisiae*
DATE 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 21-Nov-1997

ACCESSIONS S55093
REFERENCE S55089

```

#authors      Dedman, K.; Brown, D.; Bowman, S.
#submission   submitted to the EMBL Data Library, June 1995
#accession     S55093
##molecule_type DNA
##residues     1-475 #label DED
##cross-references EMBL:249809; NID:g854459; PID:g854463; MIPS:YMR211w
#experimental_source strain AB972
GENETICS
#map_position  13R
SUMMARY        #length 475 #molecular-weight 55312 #checksum 2719

Query Match    65.3%; Score 49; DB 2; Length 475;
Best Local Similarity 60.0%; Pred. No. 2.52e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 321 YLTRAITUGY 330
|||:|:|
QY 3 YLSTASSLEY 12

ENTRY 14
TITLE      ACBAE #type complete
ORGANISM   glutamate receptor precursor - great pond snail
DATE       31-Mar-1992 #sequence_revision 28-Oct-1994 #text_change
05-Sep-1997
ACCESSIONS S18443; S15681
REFERENCE   S18443
#authors    Hutton, M.L.; Harvey, R.J.; Barnard, E.A.; Darlison, M.G.
#journal     FEBS Lett. (1991) 292:111-114
#title       Cloning of a cDNA that encodes an invertebrate glutamate
receptor subunit.
#cross-references MUID:92070466
#accession   S18443
#status      nucleic acid sequence not shown
##molecule_type mRNA
##residues   1-917 #label HUT
##cross-references EMBL:X60086
REFERENCE   S15681
#authors    Hutton, M.L.; Bhandal, N.S.; Harvey, R.J.; Usherwood, P.N.R.;
Darlison, M.G.
#submission submitted to the EMBL Data Library, June 1991
#description PCR-mediated cloning of a cDNA that encodes a functional
molluscan glutamate receptor subunit.
#accession   S15681
##molecule_type mRNA
##residues   1-362, 'K', 364-776, 'S', 778-845, 'R', 847-886, 'S', 888-917
##label HUT
##cross-references EMBL:X60086; NID:g9628; PID:g9629
CLASSIFICATION #superfamily glutamate receptor; glutamate receptor homology
glycoprotein; ion channel; neurotransmitter receptor;
transmembrane protein
FEATURES
1-19          #domain signal sequence #status predicted #label SIG
20-917        #product glutamate receptor #status predicted #label
MAY\
429-853        #domain glutamate receptor homology #label GRN\
559-578        #domain transmembrane #status predicted #label TM1\
599-617        #domain transmembrane #status predicted #label TM2\
628-646        #domain transmembrane #status predicted #label TM3\
819-839        #domain transmembrane #status predicted #label TM4\
62,95,121,125,229,
251,261,272,418,
419,424,491,881 #binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY        #length 917 #molecular-weight 103106 #checksum 8952

Query Match    65.3%; Score 49; DB 1; Length 917;
Best Local Similarity 36.4%; Pred. No. 2.52e+01;
Matches 4; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 729 AYLTSTIDY 739
|||:|:|:|

```

```

QY 2 YLSTASSLEY 12

RESULT 15
ENTRY    T03030 #type fragment
TITLE     hypothetical protein KIAA0365 - human (fragment)
ORGANISM  #formal_name Homo sapiens #common_name man
DATE      23-Mar-1999 #sequence_revision 23-Mar-1999 #text_change
23-Mar-1999
ACCESSIONS T03030
REFERENCE   Z14651
#authors    Lamerdin, J.E.; McCreedy, P.M.; Skowronski, E.; Adamson,
A.W.; Burkhart-Schultz, K.; Gordon, L.; Kyle, A.; Ramirez,
M.; Stiagen, S.; Phan, H.; Velasco, N.; Garne, J.;
Dangnan, L.; Poundstone, P.; Christensen, M.; Georgescu,
A.; Avila, J.; Liu, S.; Altix, C.; Andreise, T.; Trankneim,
M.; Amico-Keller, G.; Coefield, J.; Duarte, S.; Lucas, S.;
Bruce, R.; Thomas, P.; Quan, G.; Kronmiller, B.; Arellano,
A.; Montgomery, M.; Ow, D.; Nolan, M.
#submission submitted to the EMBL Data Library, March 1998
#description Sequence analysis of an ~1 Mb region containing the MEF2B
gene in 19p12.
#accession   T03030
#status      preliminary
##residues   1-949 #label IAM
##cross-references EMBL:AC004447; NID:g2978446; PID:g2978447
#length 949 #checksum 5364

Query Match    65.3%; Score 49; DB 3; Length 949;
Best Local Similarity 60.0%; Pred. No. 2.52e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 475 FLSDENSLEY 484
|||:|:|:|
QY 3 YLSTASSLEY 12

```

Search completed: Thu Sep 2 12:50:54 1999
Job time : 14 secs.

THIS PAGE BLANK (USPTO)

MUSE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
on: Thu Sep 2 12:52:29 1999; Maspar time 1.40 Seconds
86.863 Million cell updates/sec
tabular output not generated.

Title: >US-08-599-226-34
Description: (1-12) from US0859226.pep
Perfect Score: 75
Sequence: 1 ASYSTASLEY 12

Scoring table: PAM 150
GAP 15

Searched: 106580 seqs, 10152877 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1T5A_COMB 2:5B_COMB 3:PCT9_COMB 4:backfiles1

Statistics: Mean 16.154; Variance 53.667; scale 0.301

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Sult	Score	Query	Match	Length	ID	Description	Pred. No.
No.							
1	50	66.7	37	3	PCT-US96-0	Sequence 5, Applicatio	7.21e+01
2	50	66.7	117	3	PCT-US96-0	Sequence 11, Applicati	7.21e+01
3	50	66.7	158	2	US-08-828-	Sequence 3, Applicatio	7.21e+01
4	50	66.7	159	2	US-08-828-	Sequence 1, Applicatio	7.21e+01
5	49	65.3	426	3	PCT-US95-1	Sequence 2, Applicatio	9.07e+01
6	49	65.3	426	1	US-08-336-	Sequence 2, Applicatio	9.07e+01
7	48	64.0	23	2	US-08-985-	Sequence 11, Applicati	1.14e+02
8	48	64.0	290	2	PCT-US95-1	Sequence 27, Applicati	1.14e+02
9	48	64.0	290	2	US-08-420-	Sequence 27, Applicati	1.14e+02
10	48	64.0	445	2	US-08-985-	Sequence 2, Applicatio	1.14e+02
11	48	64.0	1026	2	PCT-US95-0	Sequence 95, Applicati	1.14e+02
12	48	64.0	1026	2	US-08-268-	Sequence 95, Applicati	1.14e+02
13	48	64.0	1026	1	US-07-998-	Sequence 95, Applicati	1.14e+02
14	48	64.0	1026	1	US-08-453-	Sequence 95, Applicati	1.14e+02
15	48	64.0	1026	1	US-08-453-	Sequence 95, Applicati	1.14e+02
16	48	64.0	1026	3	PCT-US93-1	Sequence 95, Applicati	1.14e+02
17	48	64.0	1203	3	PCT-US95-0	Sequence 103, Applicati	1.14e+02
18	48	64.0	1203	3	PCT-US93-1	Sequence 103, Applicati	1.14e+02
19	48	64.0	1203	1	US-07-998-	Sequence 103, Applicati	1.14e+02
20	48	64.0	1203	2	US-08-268-	Sequence 103, Applicati	1.14e+02
21	48	64.0	1203	1	US-08-453-	Sequence 103, Applicati	1.14e+02
22	48	64.0	1203	1	US-08-453-	Sequence 103, Applicati	1.14e+02
23	48	64.0	1255	2	US-08-468-	Sequence 68, Applicati	1.14e+02

RESULT	1	STANDARD:	PRT:	37	AA.
ID	PCT-US96-08730-5				
AC	xxxxxx				
XX					
DT					
XX					
DE	Sequence 5, Application PC/TUS9608730				
XX					
CC	Sequence 5, Application PC/TUS9608730				
CC	GENERAL INFORMATION:				
CC	APPLICANT: Cassels, Frederick				
CC	APPLICANT: Anderson, Jeffrey				
CC	TITLE OF INVENTION: Methods of Raising Antibodies Against E.				
CC	TITLE OF INVENTION: Coll of the Family CSF-CFA./1				
CC	NUMBER OF SEQUENCES: 15				
CC	CORRESPONDENCE ADDRESS:				
CC	ADDRESSEE: Glenna Hendricks				
CC	STREET: P.O. Box 2509				
CC	CITY: Fairfax				
CC	STATE: VA				
CC	COUNTRY: USA				
CC	ZIP: 22031				
CC	COMPUTER READABLE FORM:				
CC	MEDIUM TYPE: Floppy disk				
CC	COMPUTER: IBM PC compatible				
CC	OPERATING SYSTEM: PC-DOS/MS-DOS				
CC	SOFTWARE: Patent Release #1.0, Version #1.25				
CC	CURRENT APPLICATION DATA:				
CC	APPLICATION NUMBER: PCT/US96/08730				
CC	FILING DATE: 03-JUN-1996				
CC	CLASSIFICATION:				
CC	ATTORNEY/AGENT INFORMATION:				
CC	NAME: Hendricks, Glenna				
CC	REGISTRATION NUMBER: 32,535				
CC	REFERENCE/DOCKET NUMBER: PCT/US96/08730				
CC	TELECOMMUNICATION INFORMATION:				
CC	TELEPHONE: (703) 591-4470				
CC	TELEFAX: (703) 591-4428				
CC	INFORMATION FOR SEQ. ID NO: 5:				
CC	SEQUENCE CHARACTERISTICS:				
CC	LENGTH: 37 amino acids				
CC	TYPE: amino acid				
CC	STRANDEDNESS: single				
CC	TOPOLOGY: unknown				
CC	MOLECULE TYPE: peptide				

CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC FRAGMENT TYPE: Internal
SQ SEQUENCE 37 AA; 3864 MW; 7776 CN;
Query Match 66.7%; Score 50; DB 3; Length 37;
Best Local Similarity 58.3%; Pred. No. 7.21e+01;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Db 22 GSYLPTAVELTY 33
:||||:| | |
QY 1 ASYLSTASSLEY 12
RESULT 2
ID PCT-US96-08730-11 STANDARD; PRT; 117 AA.
AC xxxxxx
XX
DT
Sequence 11, Application PC/TUS9608730
GENERAL INFORMATION:
APPLICANT: Cassels, Frederick
APPLICANT: Anderson, Jeffrey
APPLICANT: Carter, John Mark
TITLE OF INVENTION: Methods of Raising Antibodies Against E.
COLI OF INVENTION: Coli of the Family CSF-CRA./1
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Glenna Hendricks
STREET: P.O. Box 2509
CITY: Fairfax
STATE: VA
COUNTRY: USA
ZIP: 22031
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/08730
FILING DATE: 03-JUN-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Hendricks, Glenna
REGISTRATION NUMBER: 32,535
REFERENCE/DOCKET NUMBER: PCT/US96/08730
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 591-4470
TELEFAX: (703) 591-4428
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
SQ SEQUENCE 117 AA; 12389 MW; 76297 CN;
Query Match 66.7%; Score 50; DB 3; Length 117;
Best Local Similarity 58.3%; Pred. No. 7.21e+01;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Db 22 GSYLPTAVELTY 33
:||||:| | |
QY 1 ASYLSTASSLEY 12

RESULT 3
ID US-08-828-832-3 STANDARD; PRT; 158 AA.
AC xxxxxx
XX
DT
Sequence 3, Application US/08828832
GENERAL INFORMATION:
PATENT NO. 5827711
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purni
TITLE OF INVENTION: NOVEL SUCCINATE DEHYDROGENASE SUBUNIT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incycle Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/828,832
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0250 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1575011
SQ SEQUENCE 158 AA; 17096 MW; 135076 CN;
Query Match 66.7%; Score 50; DB 2; Length 158;
Best Local Similarity 41.7%; Pred. No. 7.21e+01;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
Db 81 AAYLNPCSAMDY 92
:||||:| | |
QY 1 ASYLSTASSLEY 12
RESULT 4
ID US-08-828-832-1 STANDARD; PRT; 159 AA.
AC xxxxxx
XX
DT
Sequence 1, Application US/08828832


```

CC Sequence 1, Application US/08828832
CC Patent No. 5827711
CC GENERAL INFORMATION:
CC APPLICANT: Lal, Preeti
CC APPLICANT: Shah, Purvi
CC TITLE OF INVENTION: NOVEL SUCCINATE DEHYDROGENASE SUBUNIT
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC City: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FASTSEQ for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/828,832
CC FILING DATE: Herewith
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: PF-0250 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 159 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: Consensus
CC CLONE: 2454416
CC SQ SEQUENCE 159 AA; 17043 MW; 129477 CN;

Query Match          66.7%; Score 50; DB 2; Length 159;
Best Local Similarity 41.7%; Pred. No. 7.21e+01;
Matches      5; Conservative      5; Mismatches    2; Indels    0; Gaps    0;

Db           82 AAYLNPCSAWDY 93
|::|:|::|
Oy          1 ASYLSTASSLEY 12

RESULT      5
ID          PCT-US95-13795-2        STANDARD:      PRT:      426 AA.
XX          xxxxxx

Sequence 2, Application PC/TUS9513795
DE          Sequence 2, Application PC/TUS9513795
CC          GENERAL INFORMATION:
CC          APPLICANT: HOLLIS, GREGORY F.
CC          APPLICANT: PATEL, MAYUR D.
CC          TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULINS
CC          NUMBER OF SEQUENCES: 4
CC          CORRESPONDENCE ADDRESS:
CC          ADDRESSEE: CHRISTINE E. CARTY
CC          STREET: 126 E. LINCOLN AVENUE; P.O. BOX 2000

```

CC	CITY:	RAHWAY
CC	STATE:	NEW JERSEY
CC	COUNTRY:	USA
CC	ZIP:	07065-0907
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE:	Floppy disk
CC	COMPUTER:	IBM PC compatible
CC	OPERATING SYSTEM:	PC-DOS/MS-DOS
CC	SOFTWARE:	PatentIn Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER:	PCT/US95/13795
CC	FILING DATE:	
CC	CLASSIFICATION:	
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME:	CARTY, CHRISTINE E.
CC	REGISTRATION NUMBER:	36,099
CC	REFERENCE/DOCKET NUMBER:	19211Y
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE:	(908) 594-6734
CC	TELEFAX:	(908) 594-4720
CC	INFORMATION FOR SEQ ID NO:	2:
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH:	426 amino acids
CC	TYPE:	amino acid
CC	STRANDEDNESS:	single
CC	TOPOLOGY:	linear
CC	MOLECULE TYPE:	protein
SQ	SEQUENCE	426 AA; 47234 MM; 1032622 CN;
Query Match		
	Best Local Similarity	65.3%; SCORE 49; DB 3; Length 426;
	Matches	6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db	214 TSYLSPSPD 224	
Oy	1 ASYSTASSLE 11	
RESULT	6	
ID	US-08-336-583-2	STANDARD; PRT; 426 AA.
XX	xxxxxx	
XX	D7	
XX	Sequence 2, Application US/08336583	
DE	Sequence 2, Application US/08336583	
CC	Patent No. 5629415	
CC	GENERAL INFORMATION:	
CC	APPLICANT:	HOLLIS, GREGORY F.
CC	APPLICANT:	PATEL, MAYUR D.
CC	TITLE OF INVENTION:	DNA ENCODING CANINE IMMUNOGLOBULIN E
CC	NUMBER OF SEQUENCES:	2
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE:	CHRISTINE E. CARTY
CC	STREET:	126 E. LINCOLN AVENUE
CC	CITY:	RAHWAY
CC	STATE:	NEW JERSEY
CC	COUNTRY:	USA
CC	ZIP:	07065-0900
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE:	Floppy disk
CC	COMPUTER:	IBM PC compatible
CC	OPERATING SYSTEM:	PC-DOS/MS-DOS
CC	SOFTWARE:	PatentIn Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER:	US/08/336,583
CC	FILING DATE:	09-NOV-1994
CC	CLASSIFICATION:	424
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME:	CARTY, CHRISTINE E.
CC	REGISTRATION NUMBER:	36,099

CC REFERENCE/DOCKET NUMBER: 19211
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (908) 594-6734
CC TELEFAX: (908) 594-4720
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 426 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 426 AA; 47234 MW; 1032622 CN;

Query Match 65.3%; Score 49; DB 1; Length 426;
Best Local Similarity 54.5%; Pred. No. 9.07e+01;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 214 TSYLSPSPDL 224
QY 1 ASYLSTASSLE 11
:|||||:|:|:
ID US-08-985-090-11 STANDARD; PRT; 23 AA.
AC xxxxxx
DT
XX
DE Sequence 11, Application US/08985090
XX
CC Sequence 11, Application US/08985090
CC Patent No. 5885893
CC GENERAL INFORMATION:
CC APPLICANT: Andrew D.J. Goodearl
CC TITLE OF INVENTION: MUSCARINIC RECEPTORS AND USES THEREFOR
CC NUMBER OF SEQUENCES: 28
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD, LLP
CC STREET: 28 State Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/985,090
CC FILING DATE:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jean M. Silveri
CC REGISTRATION NUMBER: 39,030
CC REFERENCE/DOCKET NUMBER: MNI-032
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)742-4214
CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 23 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 23 AA; 2719 MW; 3615 CN;

Query Match 64.0%; Score 48; DB 2; Length 23;
Best Local Similarity 60.0%; Pred. No. 1.14e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 2 FLITASTLEF 11
QY 3 YLSTASSLEY 12
:|||||:
ID PCT-US95-10194-27 STANDARD; PRT; 290 AA.
AC xxxxxx
DT
XX
DE Sequence 27, Application PC/TUS9510194
XX
CC Sequence 27, Application PC/TUS9510194
CC GENERAL INFORMATION:
CC APPLICANT: The Trustees of Columbia University in the City of New York
CC APPLICANT: City
CC TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
CC TITLE OF INVENTION: SEQUENCES AND USES THEREOF
CC NUMBER OF SEQUENCES: 45
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Cooper & Dunham LLP
CC STREET: 1185 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/10194
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: White, John P.
CC REGISTRATION NUMBER: 28,678
CC REFERENCE/DOCKET NUMBER: 45185-C-PCT/JPW/MSC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 278-0400
CC TELEFAX: (212) 391-0525
CC INFORMATION FOR SEQ ID NO: 27:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 290 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 290 AA; 32186 MW; 443934 CN;

Query Match 64.0%; Score 48; DB 3; Length 290;
Best Local Similarity 54.5%; Pred. No. 1.14e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 100 PYLWTPSSIIF 110
QY 2 SYLSTASSLEY 12
:|||||:
ID US-08-420-235B-27 STANDARD; PRT; 290 AA.
AC xxxxxx
DT
XX
DE Sequence 27, Application US/08420235B
XX
CC Sequence 27, Application US/08420235B
CC Patent No. 5801042

CC GENERAL INFORMATION:
CC APPLICANT: Chang, Yuan
CC APPLICANT: Moore, Patrick S.
CC TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
CC TITLE OF INVENTION: SEQUENCES AND USES THEREOF
CC NUMBER OF SEQUENCES: 47
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Cooper & Dunham LLP
CC STREET: 1185 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/420,235B
CC FILING DATE:
CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: White, John P.
CC REGISTRATION NUMBER: 28,678
CC REFERENCE/DOCKET NUMBER: 45185-B
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 278-0400
CC TELEFAX: (212) 391-0525
CC INFORMATION FOR SEQ ID NO: 27:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 290 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 290 AA; 32186 MW; 443934 CN;
SQ

Query Match 64.0%; Score 48; DB 2; Length 290;
Best Local Similarity 54.5%; Pred. No. 1.14e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 100 PVLTPSSIEF 110
:|1|1|1|:
QY 2 SYLSTASLEY 12

RESULT 10
US-08-985-090-2 STANDARD: PRT; 445 AA.
AC xxxxxx
XX
DE Sequence 2, Application US/08985090
XX
CC Sequence 2, Application US/08985090
CC Patent No. 585893
CC GENERAL INFORMATION:
CC APPLICANT: Andrew D.J. Goodearl
CC TITLE OF INVENTION: MUSCARINIC RECEPTORS AND USES THEREFOR
CC NUMBER OF SEQUENCES: 28
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD, LLP
CC STREET: 28 State Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/985,090
CC FILING DATE:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jean M. Silverl
CC REGISTRATION NUMBER: 39,030
CC REFERENCE/DOCKET NUMBER: MNI-032
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)742-4214
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 445 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 445 AA; 48671 MW; 1062278 CN;
SQ

Query Match 64.0%; Score 48; DB 2; Length 445;
Best Local Similarity 60.0%; Pred. No. 1.14e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 198 FLITASTIEF 207
:|1|1|1|1|:
QY 3 YLSTASLEY 12

RESULT 11
PCT-US95-08071-95 STANDARD: PRT; 1026 AA.
AC xxxxxx
XX
DE Sequence 95, Application PC/TUS9508071
XX
CC Sequence 95, Application PC/TUS9508071
CC GENERAL INFORMATION:
CC APPLICANT: Suzuki, Shintaro
CC TITLE OF INVENTION: Protocadherin Materials and Methods
CC NUMBER OF SEQUENCES: 115
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
CC STREET: 6300 Sears Tower, 233 S. Wacker Drive
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/08071
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/12588
CC FILING DATE: 23 DEC 1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/998,003
CC FILING DATE: 29 DEC 1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Noland, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 32149
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/474-6300

CC TELEFAX: 312/474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1026 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1026 AA; 111270 MW; 5611711 CN;
Query Match 64.0%; Score 48; DB 3; Length 1026;
Best Local Similarity 40.0%; Pred. No. 1.14e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Db 423 FLOTTPLDY 432
QY 3 YLSTASSLEY 12
RESULT 12
US-08-268-161A-95 STANDARD; PRT; 1026 AA.
Sequence 95, Application US/08268161A
Patent No. 5798224
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
ADDRESS: Borun
STREET: 233 South Wacker, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/268,161A
FILING DATE: June 27, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Young J. Suh
REGISTRATION NUMBER: P-41,337
REFERENCE/DOCKET NUMBER: 27866/32149
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 1026 AA; 111270 MW; 5611711 CN;
Query Match 64.0%; Score 48; DB 2; Length 1026;
Best Local Similarity 40.0%; Pred. No. 1.14e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Db 423 FLOTTPLDY 432
QY 3 YLSTASSLEY 12

QY 3 YLSTASSLEY 12
RESULT 13
US-07-998-003A-95 STANDARD; PRT; 1026 AA.
Sequence 95, Application US/07998003A
Patent No. 5643781
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
ADDRESS: Bicknell
STREET: 20 South Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/998,003A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5643781and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 30903
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/346-5750
TELEFAX: 312/984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 1026 AA; 111270 MW; 5611711 CN;
Query Match 64.0%; Score 48; DB 1; Length 1026;
Best Local Similarity 40.0%; Pred. No. 1.14e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Db 423 FLOTTPLDY 432
QY 3 YLSTASSLEY 12
RESULT 14
US-08-453-695A-95 STANDARD; PRT; 1026 AA.
Sequence 95, Application US/08453695A
Patent No. 5708143
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro

CC TITLE OF INVENTION: Protocadherin Materials and Methods
CC NUMBER OF SEQUENCES: 115
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
CC ADDRESSEE: Borun
CC STREET: 233 South Wacker, 6300 Sears Tower
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/453,695A
CC FILING DATE:
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: No. 5708143and, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 32658
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/474-6300
CC TELEFAX: 312/474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1026 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1026 AA: 111270 MW: 5611711 CN;

Db 423 FLOTTPLDY 432
:| |::| |
QY 3 YLSTASSLEY 12

Query Match 64.0%; Score 48; DB 1; Length 1026;
Best Local Similarity 40.0%; Pred. No. 1.14e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 423 FLOTTPLDY 432
:| |::| |
QY 3 YLSTASSLEY 12

RESULT 15
ID US-08-453-274B-95 STANDARD; PRT: 1026 AA.
XX
DY
XXXXXX
Sequence 95, Application US/08453274B
Patent No. 5663300
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,274B

CC FILING DATE: 30-MAY-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: No. 5663300and, Greta E.
CC REGISTRATION NUMBER: 35,302
CC REFERENCE/DOCKET NUMBER: 32660
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312/474-6300
CC TELEFAX: 312/474-0448
CC TELEX: 25-3856
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1026 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1026 AA: 111270 MW: 5611711 CN;

Db 423 FLOTTPLDY 432
:| |::| |
QY 3 YLSTASSLEY 12

Query Match 64.0%; Score 48; DB 1; Length 1026;
Best Local Similarity 40.0%; Pred. No. 1.14e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Search completed: Thu Sep 2 12:52:37 1999
Job time : 8 secs.

THIS PAGE BLANK (USPTO)

CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

Query Match 100.0%; Score 75; DB 27; Length 12;
Best Local Similarity 100.0%; Pred. No. 4,20e-01;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 asystassley 12
Oy 1 ASYSTASSLEY 12

RESULT 2
W27588 standard; peptide: 12 AA.

AC W27588:
DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
PD 14-AUG-1997:
PF 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226.
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRM, Kaymakalan Z, Labkovsky B,
PI Markovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
PI WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
PS Claim 20, page 73; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

Query Match 96.0%; Score 72; DB 27; Length 12;
Best Local Similarity 83.3%; Pred. No. 8,89e-01;

Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 1 asystassley 12
Oy 1 ASYSTASSLEY 12

RESULT 3
W27569 standard; Protein: 121 AA.

AC W27569:
DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain variable region.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody;
KW heavy chain; variable region; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
PD 14-AUG-1997:
PF 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226.
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRM, Kaymakalan Z, Labkovsky B,
PI Markovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
PI WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
PS Claim 16; page 76; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain variable region.
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 121 AA;

Query Match 93.3%; Score 70; DB 27; Length 121;
Best Local Similarity 90.9%; Pred. No. 1,46e+00;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 100 sylstassley 110
Oy 2 SYLSTASSLEY 12

RESULT 4
W27592 standard; peptide: 12 AA.

AC W27592:
DE 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;

KM heavy chain: complementarity determining region 3; inhibition;
 treatment; sepsis; disease; autoimmune disease; infectious disease;
 KM malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KM keloid formation; scar tissue formation; pyrexia; HUVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WPI: 97-415302/38.
 CC High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:
 SQ

Query Match 90.7%; Score 68; DB 27; Length 12;
 Best Local Similarity 83.3%; Pred. No. 2.39e+00;
 Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 DB 1 asfistssley 12
 QY 1 ASYLSTASSLEY 12

RESULT 5
 ID W2591 standard; peptide: 12 AA.
 AC W2591;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KM Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KM heavy chain: complementarity determining region 3; inhibition;
 KM treatment; sepsis; disease; autoimmune disease; infectious disease;
 KM malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KM keloid formation; scar tissue formation; pyrexia; HUVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WPI: 97-415302/38.
 CC High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 74; 102pp; English.
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:
 SQ

PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WPI: 97-415302/38.
 CC High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:
 SQ

Query Match 89.3%; Score 67; DB 27; Length 12;
 Best Local Similarity 83.3%; Pred. No. 3.06e+00;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 DB 1 asylstssley 12
 QY 1 ASYLSTASSLEY 12

RESULT 6
 ID W2590 standard; peptide: 12 AA.
 AC W2590;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KM Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KM heavy chain: complementarity determining region 3; inhibition;
 KM treatment; sepsis; disease; autoimmune disease; infectious disease;
 KM malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KM keloid formation; scar tissue formation; pyrexia; HUVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN WO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; 002219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 Manovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 WPI: 97-415302/38.
 CC High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:
 SQ

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption diseases,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SO Sequence 12 AA;

Query Match 85.3%; Score 64; DB 27; Length 12;
Best Local Similarity 75.0%; Pred. No. 6.36e+00;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 1 asy1stss1d 12
| | | | | : | | : |
Oy 1 ASYSTASSLEY 12

RESULT 7
ID W27586 standard; peptide: 12 AA.
AC W27586;
DT 19-MAR-1998 (first entry)

DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha: TNF-alpha: antibody: CDR3:
KW heavy chain: complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN MO9729131-A1.
PD 14-AUG-1997.
PE 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PS 09-FEB-1996: US-599226.

PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 72; 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,

CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SO Sequence 12 AA;

Query Match 80.0%; Score 60; DB 27; Length 12;
Best Local Similarity 81.8%; Pred. No. 1.66e+01;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 asy1stss1d 11
| | | | | : | | : |
Oy 1 ASYSTASSLEY 11

RESULT 8
ID W27589 standard; peptide: 12 AA.
AC W27589;
DT 19-MAR-1998 (first entry)

DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha: TNF-alpha: antibody: CDR3:
KW heavy chain: complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN MO9729131-A1.
PD 14-AUG-1997.
PE 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PS 09-FEB-1996: US-599226.

PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 73; 102pp: English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SO Sequence 12 AA;

Query Match 80.0%; Score 60; DB 27; Length 12;
Best Local Similarity 81.8%; Pred. No. 1.66e+01;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 asy1stss1d 11
| | | | | : | | : |
Oy 1 ASYSTASSLEY 11

PA (BADI) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Markovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Safiedd JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR WPI: 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Disclosure: Page 75; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SO Sequence 12 AA:

Query Match 77.3%; Score 58; DB 27; Length 12;
 Best Local Similarity 90.0%; Pred. No. 2.68e+01;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 2 sylstasid 11
 |||||
 Qy 2 SYLSTASSLE 11

RESULT 12
 ID W55260 standard; Protein: 159 AA.
 AC W55260:
 DT 02-JUL-1998 (first entry)
 DE H. pylori ORF 06cp30603orf16 protein.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW Identification; binding compound; bacteria; life cycle; activator;
 KW Inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 PS Helicobacter pylori.
 PS W09737044-A1.
 DR 09-OCT-1997:
 DR 27-MAR-1997: U05223.
 PR 06-DEC-1996: US-761318.
 PR 29-MAR-1996: US-625811.
 PR 02-APR-1996: US-758731.
 PR 25-OCT-1996: US-736905.
 PR 28-OCT-1996: US-738859.
 PA (ASTR) ASTRA AB.
 PI Alm RA, Smith D.
 PI WPI: 97-503122/46.
 DR N-PSDB: V24669.
 PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection
 PS Claim 14; Page 502; 1145pp; English.
 CC This sequence is a H. pylori protein of unspecified function.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds, The
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance

CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. To identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 SO Sequence 159 AA:

Query Match 76.0%; Score 57; DB 29; Length 159;
 Best Local Similarity 80.0%; Pred. No. 3.39e+01;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 60 ylsiasaley 69
 ||| |||
 Qy 3 YLSTASSLEY 12

RESULT 13
 ID W55472 standard; Protein: 177 AA.
 AC W55472:
 DT 24-JUN-1998 (first entry)
 DE H. pylori ORF 06cp30603_23452_c3.80 inner membrane protein.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW Identification; binding compound; bacteria; life cycle; activator;
 KW Inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 PS Helicobacter pylori.
 PS W09737044-A1.
 DR 09-OCT-1997:
 DR 27-MAR-1997: U05223.
 PR 06-DEC-1996: US-761318.
 PR 29-MAR-1996: US-625811.
 PR 02-APR-1996: US-758731.
 PR 25-OCT-1996: US-736905.
 PR 28-OCT-1996: US-738859.
 PA (ASTR) ASTRA AB.
 PI Alm RA, Smith D.
 PI WPI: 97-503122/46.
 DR N-PSDB: V24881.
 PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection
 PS Claim 14, 80; Page 679; 1145pp; English.
 CC This sequence is a H. pylori cell envelope inner membrane protein
 CC involved in cofactor metabolism.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds, The
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance
 CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. To identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 SO Sequence 177 AA:

Query Match 76.0%; Score 57; DB 29; Length 177;
 Best Local Similarity 80.0%; Pred. No. 3.39e+01;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 60 ylsasaley 69
 ||| |||
 QY 3 YLSTASLEY 12

RESULT 14
 ID W55688 standard; Protein: 526 AA.
 AC W55688;

DE 07-JUL-1998 (first entry)
 DE H. pylori ORF 09cpl0713_23452_c3_195 inner membrane protein.
 DE Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 DE identification; binding compound; bacteria; life cycle; activator;
 OS Helicobacter pylori.
 OS Helicobacter pylori.

PN W09737044-A1.
 PD 09-OCT-1997.
 PF 27-MAR-1997; U05223.
 PF 06-DEC-1996; US-761318.
 PF 29-MAR-1996; US-625811.
 PF 02-APR-1996; US-758731.
 PF 25-OCT-1996; US-736905.
 PR 28-OCT-1996; US-738859.
 PR (ASTR) ASTRA AB.

PI Alm RA, Smith D;
 PI WPI: 97-503122/46.
 DR N-PSDB: V25097.

PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptides) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection
 PS Claims 14,80; Pages 947-948; 1145pp; English.

CC This sequence is a H. pylori cell envelope inner membrane protein
 CC involved in cofactor metabolism.

CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds.
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance
 CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. To identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.

CC Sequence 526 AA;

Query Match 76.0%; Score 57; DB 29; Length 526;

Best Local Similarity 80.0%; Pred. No. 3.39e+01;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 60 ylsasaley 69
 ||| |||
 QY 3 YLSTASLEY 12

RESULT 15

ID W34987 standard; Protein: 875 AA.

AC W34987;

DE 21-MAY-1998 (first entry)
 DE Bankia gouldi endoglucanase.

DE Endoglucanase; cellulase; carboxymethylcellulose; cellulose;

KM biomass; beta-1,4-glycosidic bond; hydrolysis; saccharification;

KW thermostable enzyme; thermophilic; glycosidase.

OS Bankia gouldi mix (Clone 376P2).

PN W09744361-A1.

PD 27-NOV-1997.
 PF 22-MAY-1997; U08793.
 PR 22-MAY-1996; US-651572.
 PA (RECO-) RECOMBINANT BIOCATALYSIS INC.
 PI Lam DE Mathur EJ;
 DR WPI: 98-018435/02.

DR N-PSDB: T94195.

PT Endoglucanase(s), preferably from archaeal bacterium, AEP111a -
 PT useful to degrade carboxymethylcellulose and hydrolyse of
 PT beta-1,4-glycosidic bonds in cellulose

PS Claim 1, Fig 1C; 164pp; English.

CC This protein comprises an endoglucanase of Bankia gouldi mix (clone
 CC 376P2) that is capable of degrading carboxymethylcellulose and of
 CC hydrolysing the beta-1,4-glycosidic bonds in cellulose. It has
 CC homology to an endoglucanase of archaeobacterium AEP111a (see
 CC W34985). It can be produced from native cells or from recombinant
 CC host cells, especially prokaryotic host cells transformed with a
 CC plasmid or virus-derived vector including the endoglucanase DNA
 CC (see T94195). 24 Endoglucanases (see W34986-W35008) are claimed.
 CC They can be used to degrade cellulose for the conversion of plant
 CC biomass into fuels and chemicals, for use in detergents, textiles,
 CC animal feed, waste treatment, and in the fruit juice and brewing
 CC industries for the clarification and extraction of juices.

CC Sequence 875 AA;

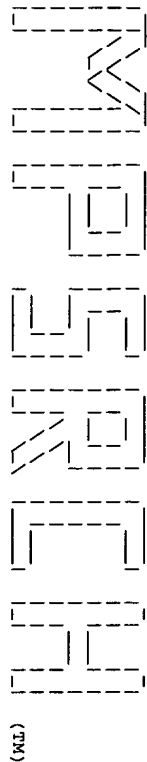
Query Match 70.7%; Score 53; DB 28; Length 875;

Best Local Similarity 58.3%; Pred. No. 8.60e+01;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 618 psylltydsly 629
 ||| |||
 QY 1 ASYLSTASLEY 12

Search completed: Thu Sep 2 12:50:23 1999
 Job time : 22 secs.

THIS PAGE BLANK (USPTO)



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Maspar.hlp protein - protein database search, using Smith-Waterman algorithm
on: Thu Sep 2 12:46:15 1999; Maspar time 3.60 Seconds
70.804 Million cell updates/sec
Abular output not generated.

Title: >US-08-599-226-33
Description: (1-12) from US08599226.pep
Perfect Score: 74
Sequence: 1 ASFLSTSSSLRY 12

Scoring table: PAM 150
Gap 15

Searched: 170751 seqs, 2126608 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneSeq35
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39

Statistics: Mean 17.466; Variance 53.940; scale 0.324
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	74	100.0	12 27	W27592	Anti-TNF-alpha antibo	3.21e+01
2	69	93.2	12 27	W27588	Anti-TNF-alpha antibo	1.19e+00
3	68	91.9	12 27	W27593	Anti-TNF-alpha antibo	1.58e+00
4	64	86.5	12 27	W27591	Anti-TNF-alpha antibo	4.31e+00
5	63	85.1	12 27	W27569	Anti-TNF-alpha antibo	5.56e+00
6	61	82.4	12 27	W27590	Anti-TNF-alpha antibo	9.22e+00
7	57	77.0	12 27	W27586	Anti-TNF-alpha antibo	2.30e+01
8	57	77.0	12 27	W27589	Anti-TNF-alpha antibo	2.50e+01
9	57	77.0	12 27	W27587	Anti-TNF-alpha antibo	2.50e+01
10	52	70.3	1026 17	R87146	Proteoglycan clone 4	8.47e+01
11	52	70.3	1026 17	R58906	Human proteoglycan	8.47e+01
12	52	70.3	1203 17	R87152	Alternatively spliced	8.47e+01
13	52	70.3	1203 17	R58911	Product of alternatively	8.47e+01
14	51	68.9	12 27	W27563	Anti-TNF-alpha antibo	1.08e+02
15	51	68.9	12 27	W27594	Anti-TNF-alpha antibo	1.08e+02
16	50	67.6	159 29	W55260	H. pylori ORF 06cp306	1.37e+02

17	50	67.6	177 29	W55472	H. pylori ORF 06cp306	1.37e+02
18	50	67.6	235 17	R91543	Pertussis toxin clone	1.37e+02
19	50	67.6	526 29	W55688	H. pylori ORF 09cp107	1.37e+02
20	49	66.2	114 12	R62919	Human cytomegalovirus	1.73e+02
21	49	66.2	509 15	R88465	Hamster scavenger rec	1.73e+02
22	48	64.9	159 37	W80222	Human succinate-ubiqu	2.19e+02
23	48	64.9	875 28	W34987	Bankia gouldi endoglu	2.19e+02
24	47	63.5	287 14	R73013	Functional equivalent	2.76e+02
25	46	62.2	218 21	R86811	Saccharomyces cerevis	3.48e+02
26	46	62.2	311 38	W89812	Protein encoded by cl	3.48e+02
27	46	62.2	411 21	R86810	Saccharomyces cerevis	3.48e+02
28	46	62.2	882 8	R43959	N-heparan sulphate su	3.48e+02
29	45	60.8	312 33	W65043	Forsythia pinoresinol	4.38e+02
30	45	60.8	312 33	W65038	Forsythia pinoresinol	4.38e+02
31	45	60.8	312 32	W65039	Forsythia pinoresinol	4.38e+02
32	45	60.8	312 32	W65042	Forsythia pinoresinol	4.38e+02
33	45	60.8	372 7	R37996	Sequence of human mac	4.38e+02
34	45	60.8	408 20	W10082	N-delta-3/C-delta-411	4.38e+02
35	45	60.8	450 38	W86335	Kidney injury associa	4.38e+02
36	45	60.8	519 20	W10072	N-delta-3/C-delta-221	4.38e+02
37	45	60.8	519 20	W10084	N-delta-3/C-delta-221	4.38e+02
38	45	60.8	520 2	R12111	Human M-CSF derivativ	4.38e+02
39	45	60.8	536 20	W10066	Human long form CSF-1	4.38e+02
40	45	60.8	536 1	P80764	Deduced amino acid se	4.38e+02
41	45	60.8	536 24	W22613	Human colony stimul	4.38e+02
42	45	60.8	554 1	P91898	Sequence of a macroph	4.38e+02
43	45	60.8	554 9	R48677	Human M-CSF	4.38e+02
44	45	60.8	554 2	P81040	Sequence of human col	4.38e+02
45	45	60.8	554 7	R37997	Sequence of human mac	4.38e+02

ALIGNMENTS

RESULT 1
ID W27592 standard; peptide: 12 AA.
AC W27592;
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disease; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HIVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN M09729131-AI.
PD 14-AUG-1997.
PE 10-FEB-1997; U02219.
PR 25-NOV-1996; US-031476.
PR 09-FEB-1996; US-599226. *W27592*
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRM, Kaymakcan Z, Labkovsky B,
PI Mankovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Sealford JG, Schoenhaut D, Vaughan TV, White M, Willon AJ;
PI WPI; 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity; e.g. to treat autoimmune diseases and cancer
PS Claim 20; Page 74; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC L29 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple

CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA:

Query Match 100.0% Score 74: DB 27: Length 12:
Best Local Similarity 100.0% Pred. No. 3,21e+01:
Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 1 asfstssley 12
||:|||||:|
Oy 1 ASFLSTSSLEY 12

RESULT 2
W27588 standard: peptide: 12 AA.

AC W27588:
DT 19-MAR-1998 (first entry)
KW Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
PD 14-AUG-1997:
PF 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226.
PA (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Markovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
TNF alpha activity, e.g. to treat autoimmune diseases and cancer
Claim 20: Page 73: 102pp; English.
The present sequence is a novel anti-human tumour necrosis
factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
determining region 3 (CDR3).
The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
less and has a Koff rate constant of 1x10 power -3 s power -1 or
less (both determined by surface plasmon resonance), and
neutralises human TNF-alpha cytotoxicity in a standard in vitro
L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
inhibits TNF-alpha activity, can be used to treat sepsis,
autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
syndrome, infectious diseases, malignancy, pulmonary, intestinal,
cardiac or inflammatory bone disorders, bone resorption disease,
alcoholic, viral or fulminant hepatitis, coagulation disturbances,
burns, reperfusion injury, keloid formation, scar tissue formation,
pyrexia, periodontal disease, obesity and radiation toxicity. The
Ab also inhibits TNF-alpha induced expression of endothelial cell
leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
endothelial cells (HUVEC).
SQ Sequence 12 AA:

Query Match 93.2% Score 69: DB 27: Length 12:
Best Local Similarity 83.3% Pred. No. 1.19e+00:

Matches 10: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

Db 1 asfstssley 12
||:|||||:|
Oy 1 ASFLSTSSLEY 12

RESULT 3
W27593 standard: peptide: 12 AA.

ID W27593:
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
PD 14-AUG-1997:
PF 10-FEB-1997: U02219.
PR 25-NOV-1996: US-031476.
PR 09-FEB-1996: US-599226.
PA (BADI) BASF AG.

PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Markovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
TNF alpha activity, e.g. to treat autoimmune diseases and cancer
Claim 20: Page 75: 102pp; English.
The present sequence is a novel anti-human tumour necrosis
factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
determining region 3 (CDR3).
The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
less and has a Koff rate constant of 1x10 power -3 s power -1 or
less (both determined by surface plasmon resonance), and
neutralises human TNF-alpha cytotoxicity in a standard in vitro
L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
inhibits TNF-alpha activity, can be used to treat sepsis,
autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
syndrome, infectious diseases, malignancy, pulmonary, intestinal,
cardiac or inflammatory bone disorders, bone resorption disease,
alcoholic, viral or fulminant hepatitis, coagulation disturbances,
burns, reperfusion injury, keloid formation, scar tissue formation,
pyrexia, periodontal disease, obesity and radiation toxicity. The
Ab also inhibits TNF-alpha induced expression of endothelial cell
leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
endothelial cells (HUVEC).
SQ Sequence 12 AA:

Query Match 91.9% Score 68: DB 27: Length 12:
Best Local Similarity 83.3% Pred. No. 1.34e+00:
Matches 10: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

Db 1 asfstssley 12
||:|||||:|
Oy 1 ASFLSTSSLEY 12

RESULT 4
W27591 standard: peptide: 12 AA.

ID W27591:
DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;

KM heavy chain; complementarity determining region 3; inhibition;
 KM treatment; sepsis; disease; autoimmune disease; infectious disease;
 KM malignancy; pulmonary disease; intestinal disorder; hepatitis;
 KM cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KM keloid formation; scar tissue formation; pyrexia; HUVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN MO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; U02219.
 PR 25-NOV-1996; US-031476.
 PS 09-FEB-1996; US-599226.
 PA (BAD) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,
 PI Mankovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 PI WPI: 97-415302/38.
 PS High affinity antibodies against human TNF alpha - useful to inhibit
 PS TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA;
 SQ

Query Match 86.5%; Score 64; DB 27; Length 12;
 Best Local Similarity 83.3%; Pred. No. 4.31e+00;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 asylstssahy 12
 11:|||||||1
 1 ASFLSTSSSLSEY 12

RESULT 5
 ID W27569 standard; Protein; 121 AA.
 AC W27569;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain variable region.
 KM Human; tumour necrosis factor-alpha; TNF-alpha; antibody;
 KM heavy chain; variable region; inhibition;
 KM treatment; sepsis; disease; autoimmune disease; infectious disease;
 KM malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KM keloid formation; scar tissue formation; pyrexia; HUVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN MO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; U02219.

PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BAD) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,
 PI Mankovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 PI WPI: 97-415302/38.
 DR N-PSDB: T88404.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 16; Page 76; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain variable region.
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 121 AA;
 SQ

Query Match 85.1%; Score 63; DB 27; Length 121;
 Best Local Similarity 72.7%; Pred. No. 5.56e+00;
 Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 DB 100 systassidy 110
 1:|||||||1
 2 SFLSTSSSLSEY 12

RESULT 6
 ID W27590 standard; peptide; 12 AA.
 AC W27590;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KM Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KM heavy chain; complementarity determining region 3; inhibition;
 KM treatment; sepsis; disease; autoimmune disease; infectious disease;
 KM malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KM cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KM bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KM keloid formation; scar tissue formation; pyrexia; HUVEC;
 KM periodontal disease; obesity; radiation toxicity;
 KM endothelial cell leukocyte adhesion molecule-1;
 KM human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN MO9729131-A1.
 PD 14-AUG-1997.
 PF 10-FEB-1997; U02219.
 PR 25-NOV-1996; US-031476.
 PS 09-FEB-1996; US-599226.
 PA (BAD) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,
 PI Mankovich JA, McGuiness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 PI WPI: 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 74; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koif rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
CC Spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

Query Match 82.4%; Score 61; DB 27; Length 12;
Best Local Similarity 75.0%; Pred. No. 9.22e+00;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Dd 1 asyistsfsldy 12
||:|||||11:1
Qy 1 ASFLSTSSSLEY 12

RESULT 7
ID W27586 standard; peptide: 12 AA.

AC W27586; DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
PD 14-AUG-1997.
PF 10-FEB-1997; U02219.
PR 25-NOV-1996; US-031476.
PR 09-FEB-1996; US-599226.
PR (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.
PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity; e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 72; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koif rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
CC Spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

Query Match 77.0%; Score 57; DB 27; Length 12;
Best Local Similarity 81.8%; Pred. No. 2.50e+01;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Dd 1 asyistsfsldy 11
||:|||||11:1
Qy 1 ASFLSTSSSLE 11

RESULT 8
ID W27589 standard; peptide: 12 AA.

AC W27589; DT 19-MAR-1998 (first entry)
DE Anti-TNF-alpha antibody heavy chain CDR3.
KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
KW heavy chain; complementarity determining region 3; inhibition;
KW treatment; sepsis; disease; autoimmune disease; infectious disease;
KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
KW keloid formation; scar tissue formation; pyrexia; HUVEC;
KW periodontal disease; obesity; radiation toxicity;
KW endothelial cell leukocyte adhesion molecule-1;
KW human umbilical vein endothelial cell.
OS Homo sapiens.
PN W09729131-A1.
PD 14-AUG-1997.
PF 10-FEB-1997; U02219.
PR 25-NOV-1996; US-031476.
PR 09-FEB-1996; US-599226.
PR (BADI) BASF AG.
PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
PI Mankovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
DR WPI: 97-415302/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
PT TNF alpha activity; e.g. to treat autoimmune diseases and cancer
PS Claim 20: Page 73; 102pp; English.
CC The present sequence is a novel anti-human tumour necrosis
CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
CC determining region 3 (CDR3).
CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
CC less and has a Koif rate constant of 1x10 power -3 s power -1 or
CC less (both determined by surface plasmon resonance), and
CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
CC inhibits TNF-alpha activity, can be used to treat sepsis,
CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
CC Spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
CC cardiac or inflammatory bone disorders, bone resorption disease,
CC burns, reperfusion injury, keloid formation, scar tissue formation,
CC pyrexia, periodontal disease, obesity and radiation toxicity. The
CC Ab also inhibits TNF-alpha induced expression of endothelial cell
CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
CC endothelial cells (HUVEC).
SQ Sequence 12 AA;

Query Match 77.0%; Score 57; DB 27; Length 12;
Best Local Similarity 81.8%; Pred. No. 2.50e+01;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Dd 1 asyistsfsldy 11
||:|||||11:1
Qy 1 ASFLSTSSSLE 11

RESULT 9
 ID W27587 standard; peptide: 12 AA.
 AC W27587;
 DT 19-MAR-1998 (first entry)
 DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarity determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HUVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.
 OS Homo sapiens.
 PN W09729131-A1.
 PR 14-AUG-1997.
 PR 10-FEB-1997; U02219.
 PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PR (BAD1) BASF AG.
 PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,
 PI Mankevich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfeld JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR WPI: 97-415302/38.
 PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 20; Page 73; 102pp; English.
 CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC 1929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. rheumatoid arthritis, allergy, multiple
 CC spondylitis, osteoarthritis, gouty arthritis, rheuma, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption diseases,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,
 CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 CC Sequence 12 AA:
 SQ

Query Match 77.0%; Score 57; DB 27; Length 12;
 Best Local Similarity 81.8%; Pred. No. 2.50e+01;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1 asystsssid 11
 11:|||||11:
 QY 1 ASFLSTSSLE 11

RESULT 10
 ID R87146 standard; Protein: 1026 AA.
 AC R87146;
 DT 29-AUG-1996 (first entry)
 DE Protocadherin clone 42.
 KW Protocadherin; pc3; pc4; pc5; human; rat; cadherin; cell adhesion; mouse;
 KW catenin; therapy; clone; frog; fruit fly.
 OS Homo sapiens.
 PN W09600289-A1.
 PD 04-JAN-1996.
 PF 26-JUN-1995; U08071.

PR 27-JUN-1994; US-268161.
 PA (DOHE-) DOHEENT EYE INST.
 PI Suzuki S;
 DR WPI: 96-068873/07.
 DR N-PSDB; T03621.
 PT Polynucleotide(s) encoding human protocadherins pc3 and pc4 and rat
 PT pc5 - involved in cell-cell adhesion and regulation activities
 PS Example 3; Page 71-75; 146pp; English.
 CC This sequence represents a clone of the human protocadherin sequence,
 CC designated pc42. The cDNA encoding this sequence was isolated after
 CC screening a human foetal brain cDNA library (contained within lambdaZapII
 CC vectors), with 32P labelled versions of the sequences represented by
 CC T03605 and T03606. The cytoplasmic domain of cadherin interacts with the
 CC cytoskeleton through catenins and other cytoskeleton associated proteins.
 CC The cytoplasmic domain is not present in all cadherins, but in those
 CC which possess it, it is essential for the cadherins adhesive function.
 CC The cadherins which do not possess a cytoplasmic domain appear to
 CC function via a different method from those with a cytoplasmic domain.
 CC This protein sequence is involved in cell-cell adhesion. This sequence
 CC may have regulatory functions in the cell, as well as the cell-cell
 CC adhesive properties. Antibodies produced against this sequence are
 CC useful for modulating the binding activity of protocadherins, and can be
 CC used therapeutically.
 CC Sequence 1026 AA:
 SQ

Query Match 70.3%; Score 52; DB 17; Length 1026;
 Best Local Similarity 50.0%; Pred. No. 8.47e+01;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 423 flgtitpidly 432
 11:|||||11:
 QY 3 FLSTSSLEY 12

RESULT 11
 ID R58906 standard; Protein: 1026 AA.
 AC R58906;
 DT 17-APR-1995 (first entry)
 DE Human protocadherin-42.
 KW Cadherin; protocadherin; cell adhesion molecule.
 OS Homo sapiens.
 PN W09414960-A.
 PD 07-JUL-1994.
 PF 23-DEC-1993; U12588.
 PR 29-DEC-1992; US-998003.
 PA (DOHE-) DOHEENT EYE INST.
 PI Suzuki S;
 DR WPI: 94-293849/36.
 DR N-PSDB; 068997.
 PT Polynucleotide sequences encoding new proto:cadherins - useful
 PT for modulating natural binding and regulating activities.
 PS Claim 13; Page 66-70; 114pp; English.
 CC Two full length human cDNAs corresp. to the partial cDNAs of
 CC HUMAN-42 and HUMAN-43 (068981,068982) were isolated
 CC from human fetal brain cDNA library. Several overlapping cDNA
 CC clones were isolated with each probe including two cDNAs which
 CC contained the putative entire coding sequences of two novel
 CC proteins designated protocadherin-42 (pc42) and protocadherin-43
 CC (pc43). The DNA and deduced AA sequences of pc42 are in
 CC 068997/R58906, while those of pc43 are in 068998/R58907. The
 CC overall structures of pc42 and pc43 are similar to that of
 CC typical cadherins but they do have distinct features. Both lack
 CC the prosequences that are present in all known cadherin precursors.
 CC The extracellular domains of pc42 and pc43 are different in length
 CC and pc42 contains seven subdomains that closely resemble the
 CC typical cadherin extracellular subdomain while pc43 has six such
 CC domains. The sequences do not show any significant homology with
 CC those of known cadherins or cadherin-related proteins.
 CC Sequence 1026 AA:
 SQ

Query Match 70.3%; Score 52; DB 11; Length 1026;
 Best Local Similarity 50.0%; Pred. No. 8.47e+01;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Dh 423 flgttptldy 432
 |||::|:|
 3 FLSTSSSLEY 12

RESULT 12
 ID R87152 standard; Protein; 1203 AA.

AC R87152;
 DT 29-AUG-1996 (first entry)
 DE Alternatively spliced pc42.
 KW Protocadherin; pc3; pc4; pc5; human; rat; cadherin; cell adhesion; mouse;
 KM catenin; therapy; clone; frog; fruit fly.
 OS Homo sapiens.
 PN WO9600289-A1.
 PD 04-JAN-1996.
 PE 26-JUN-1995; U08071.
 PR 27-JUN-1994; US-268161.
 PA (DOHE-) DOHENY EYE INST.
 PI Suzuki S.
 WPI: 96-068873/07.

N-PSDB: T03623.
 PT Polynucleotide(s) encoding human protocadherins pc3 and pc4 and rat

PS pc5 - involved in cell-cell adhesion and regulation activities
 PS Example 3; Page 90-96; 146pp: English.

CC This sequence represents a possible alternatively spliced version of the
 CC clone of the human protocadherin sequence, designated pc42. The cDNA
 CC encoding this sequence was isolated after screening a human foetal brain
 CC cDNA library (contained within lambdaZapII vectors), with 32P labelled
 CC versions of the sequences represented by T03605 and T03606. The most
 CC abundant spliced version of pc42 is represented in T03621. The
 CC cytoplasmic domain of cadherin interacts with the cytoskeleton through
 CC catenins and other cytoskeleton associated proteins. The cytoplasmic
 CC domain is not present in all cadherins, but in those which possess it, it
 CC is essential for the cadherin adhesive function. The cadherins which do
 CC not possess a cytoplasmic domain appear to function via a different
 CC method from those with a cytoplasmic domain. This protein sequence is
 CC involved in cell-cell adhesion. This sequence may have regulatory
 CC functions in the cell, as well as the cell-cell adhesive properties.
 CC Antibodies produced against this sequence are useful for modulating the
 CC binding activity of protocadherins, and can be used therapeutically.
 SQ Sequence 1203 AA;

Query Match 70.3%; Score 52; DB 17; Length 1203;
 Best Local Similarity 50.0%; Pred. No. 8.47e+01;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Dh 423 flgttptldy 432
 |||::|:|
 3 FLSTSSSLEY 12

RESULT 13
 ID R58911 standard; Protein; 1203 AA.

AC R58911;
 DT 17-APR-1995 (first entry)
 DE Product of alternative splicing of human protocadherin-42 mRNA.
 KM Cadherin; protocadherin; cell adhesion molecule.

OS Homo sapiens.
 PN WO9414960-A.
 PD 07-JUL-1994.
 PE 23-DEC-1993; U12588.
 PR 29-DEC-1992; US-998003.
 PA (DOHE-) DOHENY EYE INST.
 PI Suzuki S.

WPI: 94-293869/36.
 DR N-PSDB: Q68999.

PT Polynucleotide sequences encoding new proto:cadherins - useful
 PT for modulating natural binding and regulating activities.
 PS Example; Page 84-89; 114pp: English.
 CC Two full length human cDNAs corresp. to the partial cDNAs of
 CC HUMAN-42 and HUMAN-43 (Q68981,Q68982) were isolated
 CC from human fetal brain cDNA library. Several overlapping cDNA

CC clones were isolated with each probe including two cDNAs which
 CC contained the putative entire coding sequences of two novel
 CC proteins designated protocadherin-42 (pc42) and protocadherin-43
 CC (pc43). The DNA and deduced AA sequences of pc42 are in
 CC Q68997/R58906, while those of pc43 are in Q68998/R58907. Sequence
 CC analysis of various overlapping protocadherin cDNA clones revealed
 CC that some clones cont. unique sequences at the 3' end. The
 CC sequences forming the boundaries of the 3' end regions are
 CC consistent with the consensus sequence of mRNA splicing, suggesting
 CC that these clones may corresp. to alternatively spliced mRNAs.
 CC The DNA and AA sequences of one possible product of alternative
 CC splicing of pc42 mRNA are given in Q68999/R58911. The DNA and AA
 CC sequences of two possible products of alternative splicing of pc43
 CC mRNA are respectively presented in Q69000/R58912 and Q69001/R49144.
 SQ Sequence 1203 AA;

Query Match 70.3%; Score 52; DB 11; Length 1203;
 Best Local Similarity 50.0%; Pred. No. 8.47e+01;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Dh 423 flgttptldy 432
 |||::|:|
 3 FLSTSSSLEY 12

RESULT 14
 ID W27563 standard; peptide; 12 AA.

AC W27563;
 DT 19-MAR-1998 (first entry)

DE Anti-TNF-alpha antibody heavy chain CDR3.
 KW Human; tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;
 KW heavy chain; complementarily determining region 3; inhibition;
 KW treatment; sepsis; disease; autoimmune disease; infectious disease;
 KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;
 KW cardiac disorder; inflammatory bone disorder; reperfusion injury;
 KW bone resorption disease; coagulation disturbance; burn; ELAM-1;
 KW keloid formation; scar tissue formation; pyrexia; HIVEC;
 KW periodontal disease; obesity; radiation toxicity;
 KW endothelial cell leukocyte adhesion molecule-1;
 KW human umbilical vein endothelial cell.

OS Homo sapiens.
 FH key Location/Qualifiers
 FT Misc_difference 12
 FT /label= Tyr, Asn

PD 14-AUG-1997.
 PE 10-FEB-1997; U02219.

PR 25-NOV-1996; US-031476.
 PR 09-FEB-1996; US-599226.
 PA (BADI) BASF AG.

PI Allen DJ, Hoogenboom HRJM, Kaymakalan Z, Labkovsky B,
 PI Mankovich JA, McGuinness BT, Roberts AJ, Sakorafas P,
 PI Salfield JG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;
 DR WPI: 97-413502/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit
 PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer
 PS Claim 9; Page 65; 102pp: English.

CC The present sequence is a novel anti-human tumour necrosis
 CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity
 CC determining region 3 (CDR3).
 CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or
 CC less and has a Koff rate constant of 1x10 power -3 s power -1 or
 CC less (both determined by surface plasmon resonance), and
 CC neutralises human TNF-alpha cytotoxicity in a standard in vitro
 CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which
 CC inhibits TNF-alpha activity, can be used to treat sepsis,
 CC autoimmune diseases, e.g. Rheumatoid arthritis, Rheumatoid
 CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple
 CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephrotic
 CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,
 CC cardiac or inflammatory bone disorders, bone resorption disease,
 CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,
 CC burns, reperfusion injury, keloid formation, scar tissue formation,

CC pyrexia, periodontal disease, obesity and radiation toxicity. The
 CC Ab also inhibits TNF-alpha induced expression of endothelial cell
 CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein
 CC endothelial cells (HUVEC).
 SQ Sequence 12 AA:

Query Match

68.9%; Score 51; DB 27; Length 12;

Best Local Similarity 70.0%; Pred. No. 1.08e+02;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

DB 2 systaasid 11

QY 2 SFLSTSSLE 11

RESULT 15

ID W27594 standard; peptide: 12 AA.

AC W27594;

BT 19-MAR-1998 (first entry)

TI Anti-TNF-alpha antibody heavy chain CDR3.

KW Human: tumour necrosis factor-alpha; TNF-alpha; antibody; CDR3;

KW heavy chain; complementarity determining region 3; inhibition;

KW treatment; sepsis; disease; autoimmune disease; infectious disease;

KW malignancy; pulmonary disorder; intestinal disorder; hepatitis;

KW cardiac disorder; inflammatory bone disorder; reperfusion injury;

KW bone resorption disease; coagulation disturbance; burn; ELAM-1;

KW keloid formation; scar tissue formation; pyrexia; HUVEC;

KW periodontal disease; obesity; radiation toxicity;

KW endothelial cell leukocyte adhesion molecule-1;

KW human umbilical vein endothelial cell.

OS Homo sapiens.

PN WO9729131-A1.

PD 14-AUG-1997.

PF 10-FEB-1997; U02219.

PR 25-NOV-1996; US-031476.

PR 09-FEB-1996; US-599226.

PA (BAD1) BASF AG.

PI Allen DJ, Hoogenboom HRJM, Kaymakcalan Z, Labkovsky B,

PI Mankevich JA, McGuiness BT, Roberts AJ, Sakorafas P,

PI Salfeld UG, Schoenhaut D, Vaughan TJ, White M, Wilton AJ;

DR WPI: 97-413502/38.

PT High affinity antibodies against human TNF alpha - useful to inhibit

PT TNF alpha activity, e.g. to treat autoimmune diseases and cancer

PS Disclosure: Page 75: 102pp: English.

CC The present sequence is a novel anti-human tumour necrosis

CC factor-alpha (TNF-alpha) antibody (Ab) heavy chain complementarity

CC determining region 3 (CDR3).

CC The Ab dissociates from TNF-alpha with a Kd of 1x10 power -8 M or

CC less and has a Koff rate constant of 1x10 power -3 s power -1 or

CC less (both determined by surface plasmon resonance), and

CC neutralises human TNF-alpha cytotoxicity in a standard in vitro

CC L929 assay with an IC50 of 1x10 power -7 M or less. The Ab, which

CC inhibits TNF-alpha activity, can be used to treat sepsis,

CC autoimmune diseases, e.g. rheumatoid arthritis, rheumatoid

CC spondylitis, osteoarthritis, gouty arthritis, allergy, multiple

CC sclerosis, autoimmune diabetes, autoimmune uveitis or nephritic

CC syndrome, infectious diseases, malignancy, pulmonary, intestinal,

CC cardiac or inflammatory bone disorders, bone resorption disease,

CC alcoholic, viral or fulminant hepatitis, coagulation disturbances,

CC burns, reperfusion injury, keloid formation, scar tissue formation,

CC pyrexia, periodontal disease, obesity and radiation toxicity. The

CC Ab also inhibits TNF-alpha induced expression of endothelial cell

CC leukocyte adhesion molecule-1 (ELAM-1) on human umbilical vein

CC endothelial cells (HUVEC).

SQ Sequence 12 AA:

Query Match

68.9%; Score 51; DB 27; Length 12;

Best Local Similarity 70.0%; Pred. No. 1.08e+02;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

DB 2 systaasid 11

QY 2 SFLSTSSLE 11

Search completed: Thu Sep 2 12:46:36 1999
 Job time : 21 secs.

THIS PAGE BLANK (USPTO)